

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 10:47:03 ; Search time 3260 Seconds
(without alignments)
16869.613 Million cell updates/sec

Title: US-09-813-290-1

Perfect score: 1 atggcctgagccctagctgg.....gggaggtgagccacgtag 2628

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

1	2627.6	100.0	2628	6	AX253545	Sequence
2	2627.6	100.0	3568	6	AX253549	Sequence
3	2348.6	89.4	2349	6	AX253547	Sequence
4	2348.6	89.4	4700	9	AB029496	Homo sapi
5	2177.6	82.9	4469	9	AK024425	Homo sapi
6	2148.6	81.8	2340	6	AX155170	Sequence
7	1471.8	56.0	2337	6	AX155172	Sequence
8	632.4	24.1	2898	10	AF034744	Mus muscu
9	629.6	24.0	2328	10	MMSEMH	Mus muscu
10	606.8	23.1	2219	9	HSU28369	Homo sapien
11	595.6	22.7	2825	9	BC013975	Homo sapi
12	594.8	22.6	2813	9	AF217991	Homo sapi
13	570.8	21.7	3988	9	MM293948	Mus muscu
14	570.8	21.7	4466	10	MM293947	Mus muscu
15	566.8	21.6	2615	5	AF022947	Gallus ga
16	561.6	21.4	6474	9	AB002329	Human mRN
17	561.6	20.8	2765	9	BC009113	Homo sapi
18	537.4	20.4	2872	10	MMRASEMA	Mus muscu
19	536.6	20.4	3352	9	HSU38276	Human semap
20	505.8	19.2	2319	10	MUSCOSE	Mus muscu
21	505.8	19.2	3148	5	MUSC1	Danio rer
22	499.4	19.0	2265	10	AF080091	Mus muscu
23	498.6	19.0	2489	10	BC010976	Mus muscu
24	498.6	19.0	2337	5	AF083382	Danio rer
25	497.8	18.9	3263	5	GGU02528	Gallus gall
26	492.6	18.7	1998	10	MUSSEMA111	Mus muscu
27	492	18.7	2913	10	MMRASEMD	Mus muscu
28	491.4	18.5	2319	10	RNSITICN1	R. norvegicu
29	485	18.1	2530	9	HUMSEMA	Homo sapien
30	474.8	18.1	2530	11	G31703	SW651973 Et
31	474.8	18.1	2530	6	AR040598	Sequence
32	474.8	18.1	2601	6	AR040598	Sequence
33	474.8	18.1	2601	6	AX207154	Sequence
34	474.8	18.1	2709	6	AX207154	Sequence
35	474	18.0	123943	9	AC006208	Homo sapi
36	474	18.0	214046	2	AC0094021	Homo sapi
37	467.6	17.8	2719	9	HSU33920	Human clone
38	457	17.4	2879	5	AF124485	Danio rer
39	445.6	17.0	2715	5	GGU28240	Gallus gall
40	433.8	16.5	2383	10	AF080090	Mus muscu
41	406	15.4	2477	10	MMRASEME	Mus muscu
42	390	14.8	5177	9	AB000220	Homo sapi
43	384.2	14.6	2331	6	E27343	Novel polyP
44	384.2	14.6	3880	6	E27343	Novel polyP
45	364	13.9	2336	5	AF022946	Gallus ga

ALIGNMENTS

RESULT 1	AX253545	2628 bp	DNA	linear	PAT 10-OCT-2001
LOCUS	AX253545				
DEFINITION	Sequence 1 from Patent WO0170806.				
ACCESSION	AX253545				
VERSION	AX253545.1	GI:16074048			

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

location/Qualifiers

1. 2628

/Organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT

ORIGIN

459 a

836 c

848 g

484 t

1 others

Db 1755 CAGCTAGAGATGTTCTCTGCTGCGCCCAAGGCGGGAAGACCTCGAGTGTACGGC 1814
 QY 1261 CTGTTACAGACCGTCAAGGCGGTGTTCAGAGGCTTGCCCTGTGTGTACACATGGCA 1320
 Db 1815 CTGTTACAGACCGTCAAGGCGGTGTTCAGAGGCTTGCCCTGTGTGTACACATGGCA 1874
 QY 1321 GACATCTGGAGAGTTTTCACAGGCGGCTTGCCACCGAGATGGGCTTCAGACCAATGG 1380
 Db 1875 GACATCTGGAGAGTTTTCACAGGCGGCTTGCCACCGAGATGGGCTTCAGACCAATGG 1934
 QY 1381 GGGGCTATGGGGGCAAGGCGGCTTGCCCTGCGGCTGTGCCCCAGCAAGATGACC 1440
 Db 1935 GGGGCTATGGGGGCAAGGCGGCTTGCCCTGCGGCTGTGCCCCAGCAAGATGACC 1994
 QY 1441 GCACAGCAGAGAGGCGCTTTTGGCAGACCAAGAGACTTACCGAGATGGTGTGAGTTT 1500
 Db 1995 GCACAGCAGAGAGGCGCTTTTGGCAGACCAAGAGACTTACCGAGATGGTGTGAGTTT 2054
 QY 1501 GCGGAGGCGGCGGCTTTCATGTTTCTGCTGTGCGGCTGTGACATGGCGGCTTGCTT 1560
 Db 2055 GCGGAGGCGGCGGCTTTCATGTTTCTGCTGTGCGGCTGTGACATGGCGGCTTGCTT 2114
 QY 1561 GTCAAGAGCCGCTGCGGCGGCTTACAGAGCTACAGAGATGTTGGAGAGCGGCTGGAGGAGAG 1620
 Db 2115 GTCAAGAGCCGCTGCGGCGGCTTACAGAGCTACAGAGATGTTGGAGAGCGGCTGGAGGAGAG 2174
 QY 1621 GATGGAGCTAGAGATGCTATTTCTGCGGAGCTGACTAGGAGTGTGCTCAAGATGATC 1680
 Db 2175 GATGGAGCTAGAGATGCTATTTCTGCGGAGCTGACTAGGAGTGTGCTCAAGATGATC 2234
 QY 1681 GCTCTCCAGAGAGGCGGCTCACTGACCTAGAGAGTGTGTTGGAGAGCTTCAGAGT 1740
 Db 2235 GCTCTCCAGAGAGGCGGCTCACTGACCTAGAGAGTGTGTTGGAGAGCTTCAGAGT 2294
 QY 1741 TTTAAGAGTCCCAAGACCTATCACCGGAATGAGATCTCTGCAAAAGCAATCTATAC 1800
 Db 2295 TTTAAGAGTCCCAAGACCTATCACCGGAATGAGATCTCTGCAAAAGCAATCTATAC 2354
 QY 1801 GTGGGCTCTGCGGCTGTGTGGCGGCTGAGCTGCGCTGACCAATGTGAGTTTGGAGT 1860
 Db 2355 GTGGGCTCTGCGGCTGTGTGGCGGCTGAGCTGCGCTGACCAATGTGAGTTTGGAGT 2414
 QY 1861 GCGTGTGAGAGTGTGCTGCGGCGGCGGAGCCCATATCTGCTGGAGTGTGCTCTGT 1920
 Db 2415 GCGTGTGAGAGTGTGCTGCGGCGGCGGAGCCCATATCTGCTGGAGTGTGCTCTGT 2474
 QY 1921 ACCCACTACCGCCCGAGCCTTGGCAAGCGCGGTTCCGCGGAGAGCAATCCGCGACGCG 1980
 Db 2475 ACCCACTACCGCCCGAGCCTTGGCAAGCGCGGTTCCGCGGAGAGCAATCCGCGACGCG 2534
 QY 1981 AACCTGCGGCTGAGTGTGCTGGGCGGAGAGCCAGAGCAAGAGAGGAGTGGAGTGTGGA 2040
 Db 2535 AACCTGCGGCTGAGTGTGCTGGGCGGAGAGCCAGAGCAAGAGAGGAGTGGAGTGTGGA 2594
 QY 2041 GCGACATGTTTACGCGAGAGAGCAATATGACCTTCTGGAGTGTGCTGCGCAAGTCT 2100
 Db 2595 GCGACATGTTTACGCGAGAGAGCAATATGACCTTCTGGAGTGTGCTGCGCAAGTCT 2654
 QY 2101 CCCCAAGCTGCTGTGGGCTGTGCTTTCAGAGGCGAGGAGTGTGAGGCGCTTACAGAGTGT 2160
 Db 2655 CCCCAAGCTGCTGTGGGCTGTGCTTTCAGAGGCGAGGAGTGTGAGGCGCTTACAGAGTGT 2714
 QY 2161 AAGAGGAGAGAGGAGTGTTCAGAGAGAGCGGCGGCTGTGCTGCGCAAGTGT 2220
 Db 2715 AAGAGGAGAGAGGAGTGTTCAGAGAGAGCGGCGGCTGTGCTGCGCAAGTGT 2274
 QY 2221 TTGAGTGGGAGAGCTTACACCTGTGACACTGTGAGCATGTGCTTCCCAAGTGTGCT 2280
 Db 2775 TTGAGTGGGAGAGCTTACACCTGTGACACTGTGAGCATGTGCTTCCCAAGTGTGCT 2834
 QY 2281 CCGCTGCTGT 2340
 Db 2835 CCGCTGCTGT 2894

QY 2341 AAGCAGAGAGAGGCGGCGGCGGAGGCGCTGTGCTTCCACCCCAAGGCGCTGTGAC 2400
 Db 2895 AAGCAGAGAGAGGCGGCGGCGGAGGCGCTGTGCTTCCACCCCAAGGCGCTGTGAC 2954
 QY 2401 AAGGACATCTGAGAGCTCATTTGCTGTGCGCAACCTGCGGCGGCTGTGAGTGTGAG 2460
 Db 2955 AAGGACATCTGAGAGCTCATTTGCTGTGCGCAACCTGCGGCGGCTGTGAGTGTGAG 3014
 QY 2461 CCGGT 2520
 Db 3015 CCGGT 3074
 QY 2521 AAGCAGGCGAGGCGGCGGCGGAGGCTGTGAGGCTGTGAGTGTGAGTGTGAGTGTGAG 2580
 Db 3075 AAGCAGGCGAGGCGGCGGCGGAGGCTGTGAGGCTGTGAGTGTGAGTGTGAGTGTGAG 3134
 QY 2581 GTGCAATGCGGAGAGCAATGTGAGGCGGCGGAGGCTGTGAGTGTGAGTGTGAGTGTGAG 2628
 Db 3135 GTGCAATGCGGAGAGCAATGTGAGGCGGCGGAGGCTGTGAGTGTGAGTGTGAGTGTGAG 3182

RESULT 3
 AX253547
 LOCUS AX253547
 DEFINITION Sequence 3 from Patent WO0170806.
 ACCESSION AX253547
 VERSION AX253547.1 GI:16074049
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2349)
 AUTHORS Walke,D.W., Wiljanowski,N.L., Turner,C.A., Hilbun,E., Wang,X.,
 Donoho,G. and Scoville,J.
 TITLE Human secreted proteins and polynucleotides encoding the same
 JOURNAL Patent: WO 0170806-A 3 27-SEP-2001.
 FEATURES
 Lexicon Genetics Incorporated (US)
 Location/Qualifiers
 source 1..2349

BASE COUNT 406 a 747 c 753 g 442 t 1 others
 ORIGIN
 Query Match 89.4%; Score 2348.6; DB 6; Length 2349;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 ATGGCCCCCTGCGGCTGTGCGGCTATTTGCTGCTAGGAGGCGCTCTGCTCATGGGGT 339
 Db 1 ATGGCCCCCTGCGGCTGTGCGGCTATTTGCTGCTAGGAGGCGCTCTGCTCATGGGGT 60
 QY 340 AGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 399
 Db 61 AGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 400 CTGTGTGCAACCGCTGTGCTGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 459
 Db 121 CTGTGTGCAACCGCTGTGCTGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 460 ATGTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 519
 Db 181 ATGTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 QY 520 CTGCGGCTGTGCAACCGCTGTGCTGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 579
 Db 241 CTGCGGCTGTGCAACCGCTGTGCTGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 580 CAGAGGAGAGAGT 639
 Db 301 CAGAGGAGAGAGT 360

QY 640 CGGTCCTACAGCTTCACACCGGACCCACCTGCTGAGGACTGGGCGCTTCAG 699
 Db 361 CGGGTGTACAGCTTCACACCGGACCCACCTGCTGAGGACTGGGCGCTTCAG 420
 QY 700 CCACCTGCTGCTTCACAGCTTCGCGCCACCGTGGGAGCATGTCTCCACCTGAGCT 759
 Db 421 CCACCTGCTGCTTCACAGCTTCGCGCCACCGTGGGAGCATGTCTCCACCTGAGCT 480
 QY 760 GGCAGTGTGAAAGTGGCGGGGGCGCTCCACAGAGCCAGCGCTTCCTCCAGC 819
 Db 481 GGCAGTGTGAAAGTGGCGGGGGCGCTCCACAGAGCCAGCGCTTCCTCCAGC 540
 QY 820 ACCTTCATAGAGGGGAGCTGTACAGAGGCTCTCAGTGTGACTTCTGGGGGAGAGGCG 879
 Db 541 ACCTTCATAGAGGGGAGCTGTACAGAGGCTCTCAGTGTGACTTCTGGGGGAGAGGCG 600
 QY 880 ATGATCTTCGAAAGTGAAGTCTCTGCGGCGAGCTCTGAGCTTCGAGACTCTC 939
 Db 601 ATGATCTTCGAAAGTGAAGTCTCTGCGGCGAGCTCTGAGCTTCGAGACTCTC 660
 QY 940 TTGACAGACCCCGGTTGTGATGGCCCGGATCCCTGAGAACTCTGACAGAGCAAT 999
 Db 661 TTGACAGACCCCGGTTGTGATGGCCCGGATCCCTGAGAACTCTGACAGAGCAAT 720
 QY 1000 GACAAAGTGTACTTCTTCTTCGAGAGAGGTCCTGCGGCGAGTGGGCTCGAACCAT 1059
 Db 721 GACAAAGTGTACTTCTTCTTCGAGAGAGGTCCTGCGGCGAGTGGGCTCGAACCAT 780
 QY 1060 GTACAGTGAAGCGGCTGGGCGCGCTGCTGCTGAGTGTCTGGGGGCGAGGCTGTG 1119
 Db 781 GTACAGTGAAGCGGCTGGGCGCGCTGCTGCTGAGTGTCTGGGGGCGAGGCTGTG 840
 QY 1120 GTGAAACAATGAGACACTTTCCTCAAGAGGCGAGGCTGTCTCGCTGCTCGGCGCTGT 1179
 Db 841 GTGAAACAATGAGACACTTTCCTCAAGAGGCGAGGCTGTCTCGCTGCTCGGCGCTGT 900
 QY 1180 GGTGCGAGAGCCACTTTCGACAGCTAGAGATGTCTCTGCTGTGCGCCAAAGCGCGG 1239
 Db 901 GGTGCGAGAGCCACTTTCGACAGCTAGAGATGTCTCTGCTGTGCGCCAAAGCGCGG 960
 QY 1240 AAGAGCTCGAGGTGTACCGGCTGTCAGACCGTACGCTGCTGTTCAGAGGCTTCGCT 1299
 Db 961 AAGAGCTCGAGGTGTACCGGCTGTCAGACCGTACGCTGCTGTTCAGAGGCTTCGCT 1020
 QY 1300 GTCTGTGTACCAACATGAGAGCATGTGGAGTTTCAAGGCGCTTTCGCGCCAGCA 1359
 Db 1021 GTCTGTGTACCAACATGAGAGCATGTGGAGTTTCAAGGCGCTTTCGCGCCAGCA 1080
 QY 1360 GATGGGCTCAGACCACTGGGGGCGCTTATGAGGGGCAAGGTGCCCTTCGCGCTGGC 1419
 Db 1081 GATGGGCTCAGACCACTGGGGGCGCTTATGAGGGGCAAGGTGCCCTTCGCGCTGGC 1140
 QY 1420 GTGTGCGCAGAGATGACGCGACAGCGAGGCGCTTTCGAGCAGCAACAGAGTAC 1479
 Db 1141 GTGTGCGCAGAGATGACGCGACAGCGAGGCGCTTTCGAGCAGCAACAGAGTAC 1200
 QY 1480 CCAGATGAGTGTGAGTTTCCCGAGGCGCCCTTCATGTTTGTGCGCTGTGCGGCT 1539
 Db 1201 CCAGATGAGTGTGAGTTTCCCGAGGCGCCCTTCATGTTTGTGCGCTGTGCGGCT 1260
 QY 1540 CGACATGAGGCGCTGCTGCTTCAGAGACCACTGCGCCAGCAGTACACAGATGCTG 1599
 Db 1261 CGACATGAGGCGCTGCTGCTTCAGAGACCACTGCGCCAGCAGTACACAGATGCTG 1320
 QY 1600 GTGAGCGGCTGAGGAGAGAGTGGAGTACATGCTATTTCTGGGAGTACTCA 1659
 Db 1321 GTGAGCGGCTGAGGAGAGAGTGGAGTACATGCTATTTCTGGGAGTACTCA 1380
 QY 1660 GGGTGTGTCTCAAGTCACTGCTCTCAGAGGAGGCGCTCAGCTGAACCTGAGAGTGT 1719
 Db 1381 GGGTGTGTCTCAAGTCACTGCTCTCAGAGGAGGCGCTCAGCTGAACCTGAGAGTGT 1440

QY 1720 GTTCTGAGAGACTCCAGTGTTTTAAAGTGGCAACACTATACCGAAATGAGATCTCT 1779
 Db 1441 GTTCTGAGAGACTCCAGTGTTTTAAAGTGGCAACACTATACCGAAATGAGATCTCT 1500
 QY 1780 GTCAAAAGCAAAATGCTATACGTGGGCTCTGCGTGGTGTGGCCAGCTGTGGCTCAC 1839
 Db 1501 GTCAAAAGCAAAATGCTATACGTGGGCTCTGCGTGGTGTGGCCAGCTGTGGCTCAC 1560
 QY 1840 CAATGTAGACTTACGCACTGCTGTGAGAGTGTGCTGGGCGCGGAGCCATATCT 1899
 Db 1561 CAATGTAGACTTACGCACTGCTGTGAGAGTGTGCTGGGCGCGGAGCCATATCT 1620
 QY 1900 GCTTGGATGTGCTCTCTCTTACCACTACCGCCAGCTTTCGCAAGCGCGCTTCGCG 1959
 Db 1621 GCTTGGATGTGCTCTCTCTTACCACTACCGCCAGCTTTCGCAAGCGCGCTTCGCG 1680
 QY 1960 CGGAGAGCATCCGCGACCGCAACCCCTGCTGAGTGGCTGGGCGAGCCAGAGAA 2019
 Db 1681 CGGAGAGCATCCGCGACCGCAACCCCTGCTGAGTGGCTGGGCGAGCCAGAGAA 1740
 QY 2020 GAGGAGTGGAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2079
 Db 1741 GAGGAGTGGAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
 QY 2080 CTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2139
 Db 1801 CTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
 QY 2140 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2199
 Db 1861 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
 QY 2200 CTGTTCCGAGGCTTACCGCTTTCGATGCGGCGACCTTACCTGACACTCTGAGCAT 2259
 Db 1921 CTGTTCCGAGGCTTACCGCTTTCGATGCGGCGACCTTACCTGACACTCTGAGCAT 1980
 QY 2260 GGTCTTCCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2319
 Db 1981 GGTCTTCCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 2320 AACCTGTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2379
 Db 2041 AACCTGTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
 QY 2380 ACCGACCCAGAGGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2439
 Db 2101 ACCGACCCAGAGGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
 QY 2440 CGGCTGATGATGCTGTGAGCGCTGTGAGCGGCTGTGAGCGGCTGTGAGCGGCT 2499
 Db 2161 CGGCTGATGATGCTGTGAGCGGCTGTGAGCGGCTGTGAGCGGCTGTGAGCGGCT 2220
 QY 2500 TTCCGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2559
 Db 2221 TTCCGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
 QY 2560 GGCAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2619
 Db 2281 GGCAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
 QY 2620 GCCAGCTAG 2628
 Db 2341 GCCAGCTAG 2349

RESULT 4
 AB029496 4700 bp mRNA linear PRI 07-JUL-2000
 LOCUS Homo sapiens mRNA for semaphorin sem2, complete cds.
 DEFINITION AB029496
 ACCESSION AB029496.1 GI:8978201
 VERSION semaphorin sem2.
 KEYWORDS semaphorin sem2.
 SOURCE Homo sapiens cDNA to mRNA.

ORGANISM

Homo sapiens

REFERENCE

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 4700)

AUTHORS

Miyajima, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

TITLE

Human semaphorin

Published Only in Database (2000) In press
2 (bases 1 to 4700)

REFERENCE

AUTHORS

Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,
Miyajima, N. and Saito, T.

TITLE

Direct Submission

Submitted (01-JUL-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group, Inage-ku Anagawa,
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp)
Tel: 81-43-201-3135, Fax: 81-43-251-9818

FEATURES

Location/Qualifiers

1..4700

source

/organism="Homo sapiens"

gene

/db_xref="taxon:9606"

CDS

/gene="sem2"

CDS

1..2349

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

CDS

/gene="sem2"

Homo sapiens

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 4700)

Miyajima, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Human semaphorin

Published Only in Database (2000) In press
2 (bases 1 to 4700)

Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Miyajima, N. and Saito, T.

Direct Submission

Submitted (01-JUL-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group, Inage-ku Anagawa,
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp)
Tel: 81-43-201-3135, Fax: 81-43-251-9818

Location/Qualifiers

1..4700

/organism="Homo sapiens"

/db_xref="taxon:9606"

/gene="sem2"

1..2349

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

Homo sapiens

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 4700)

Miyajima, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Human semaphorin

Published Only in Database (2000) In press
2 (bases 1 to 4700)

Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Miyajima, N. and Saito, T.

Direct Submission

Submitted (01-JUL-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group, Inage-ku Anagawa,
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp)
Tel: 81-43-201-3135, Fax: 81-43-251-9818

Location/Qualifiers

1..4700

/organism="Homo sapiens"

/db_xref="taxon:9606"

/gene="sem2"

1..2349

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

Homo sapiens

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 4700)

Miyajima, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Human semaphorin

Published Only in Database (2000) In press
2 (bases 1 to 4700)

Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,

Miyajima, N. and Saito, T.

Direct Submission

Submitted (01-JUL-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group, Inage-ku Anagawa,
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp)
Tel: 81-43-201-3135, Fax: 81-43-251-9818

Location/Qualifiers

1..4700

/organism="Homo sapiens"

/db_xref="taxon:9606"

/gene="sem2"

1..2349

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

/gene="sem2"

Db	1441	GTCTGGAGGAGCTCCAGAGGTGTTAAAGTGTGCAACACCTTACACCGAAATGAGATCTCT	1500
QY	1780	GTCAAAAGGCAAAATGCTATATACGTGGGCTCTCGGCTGTGGGTGTGGCCACAGCTGCGGCTCAC	1839
Db	1501	GTCAAAAGGCAAAATGCTATATACGTGGGCTCTCGGCTGTGGGTGTGGCCACAGCTGCGGCTCAC	1560
QY	1840	CAATGTGAGACTTACGCGACCTGCTGTGTCACAAAGTGTGCTGGCCCGGGAGCCCACTACTGT	1899
Db	1561	CAATGTGAGACTTACGCGACCTGCTGTGTCACAAAGTGTGCTGGCCCGGGAGCCCACTACTGT	1620
QY	1900	GGCTGGAGTGTGGCTCTCTGTACCCACTACCGCCCAAGCCTTGGTGGCAAGCGCGGTTCCGC	1959
Db	1621	GGCTGGAGTGTGGCTCTCTGTACCCACTACCGCCCAAGCCTTGGTGGCAAGCGCGGTTCCGC	1680
QY	1960	CGGACGACATCCGGACCGGCAACCTGCTCTGACAGTGTGCTGGGCGAGAGCCAGAGAA	2019
Db	1681	CGGACGACATCCGGACCGGCAACCTGCTCTGACAGTGTGCTGGGCGAGAGCCAGAGAA	1740
QY	2020	GAGCAGTGGGACTTGTGGCAGGACACCATGTCTACGGACGAGAGACAAATGACACTTC	2079
Db	1741	GAGCAGTGGGACTTGTGGCAGGACACCATGTCTACGGACGAGAGACAAATGACACTTC	1800
QY	2080	CTGAGAGTCTGCTGCCAAGTCTCCCAAGCTGCTGTGGCGCTGGCTCTTGCAGAGCCAGG	2139
Db	1801	CTGAGAGTCTGCTGCCAAGTCTCCCAAGCTGCTGTGGCGCTGGCTCTTGCAGAGCCAGG	1860
QY	2140	GATGAGGGGCTGTACACAGTGAAGACGAGACGAGAGACTTGTGCACACGAGAGGGGGCTG	2199
Db	1861	GATGAGGGGCTGTACACAGTGAAGACGAGACGAGAGACTTGTGCACACGAGAGGGGGCTG	1920
QY	2200	CTGTTCGCCAAGCTTAGCGCTTGGATGCGGGACCTACACTGACACTCTGTGAGCAT	2259
Db	1921	CTGTTCGCCAAGCTTAGCGCTTGGATGCGGGACCTACACTGACACTCTGTGAGCAT	1980
QY	2260	GGCTTCTCCACAGCTGTGTCCGCTGGCTCTGGTGTGATGTGTGGCTGTACAGCTGTGAC	2319
Db	1981	GGCTTCTCCACAGCTGTGTCCGCTGGCTCTGGTGTGATGTGTGGCTGTACAGCTGTGAC	2040
QY	2320	AACCTTTCCTCCGAGGCAAAAGCCAGAGAGCCCCACAGCCCGGGGAGGCGCTGCTTC	2379
Db	2041	AACCTTTCCTCCGAGGCAAAAGCCAGAGAGCCCCACAGCCCGGGGAGGCGCTGCTTC	2100
QY	2380	ACCCACCCACAGGCTGTGTACAAAGACATCTGACACTATTGGCTTGGCCAACTGTGCC	2439
Db	2101	ACCCACCCACAGGCTGTGTACAAAGACATCTGACACTATTGGCTTGGCCAACTGTGCC	2160
QY	2440	CGGATGATGATGATGTGAGCGCTGTGTGTGACAGGGGCAACAGGAATCTCAGCTGC	2499
Db	2161	CGGATGATGATGATGTGAGCGCTGTGTGTGACAGGGGCAACAGGAATCTCAGCTGC	2220
QY	2500	TTCGGAGCCGAGACCGGGGCAACAGAGGCGCAAGAGCTGGGCGAGGCTGTGACCTA	2559
Db	2221	TTCGGAGCCGAGACCGGGGCAACAGAGGCGCAAGAGCTGGGCGAGGCTGTGACCTA	2280
QY	2560	GGCAGAAGATGAAGCGCGGGTGCATCCGAGACATCGAGCGCCCGGAGGTGTG	2619
Db	2281	GGCAGAAGATGAAGCGCGGGTGCATCCGAGACATCGAGCGCCCGGAGGTGTG	2340
QY	2620	GCCAGTAG 2628	
Db	2341	GCCAGTAG 2349	

RESULT 5	AK024425	4469 bp	mRNA	linear	PRI 29-SEP-2000
LOCUS	AK024425				
DEFINITION	Homo sapiens mRNA for FLJ00014 protein, partial cds.				
ACCESSION	AK024425				
VERSION	AK024425.1	GI:10440358			
KEYWORDS	fts (full insert sequence).				
SOURCE	Homo sapiens adult spleen	cDNA	to mRNA,	clone:as00014.	
ORGANISM	Homo sapiens				

REFERENCE AUTHORS TITLE	JOURNAL, REFERENCE AUTHORS TITLE JOURNAL
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo. 1 (bases 1 to 4469)	
Ohara, O., Nagase, T., Kikuno, R. and Okumura, K. The nucleotide sequence of a long cDNA clone isolated from human spleen	
Published Only in Database (2000) In press	
2 (bases 1 to 4469)	
Ohara, O., Nagase, T., Kikuno, R. and Okumura, K. Direct Submission	
Submitted (24-AUG-2000) Gensuu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, to	

```

COMMENT      Fax:81-438-52-3914) Ministry of
              NEDO human cDNA sequencing project supported by
              international trade and industry of Japan: cDNA full insert
              sequencing: Research Association for Biotechnology; cDNA library
              construction, 3'-Rapid one pass sequencing and clone selection
              Kazusa DNA Research Institute.
FEATURES     Localized/Qualifiers
SOURCE       LOC44469
              1:transcript="Homo sapiens"

```

```

//tissue_type="spleen"
//dev_stage="adult"
//note="vector:pb1uescriptII SK plus"
1..2178
/gene="FLJ00014"
<1..2178
/gene="FLJ00014"

```

```

/product="F1J00014 protein"
/protein_id="BAB15715.1"
/db_xref="gi:10440359"
/translation="
IQAMNIDEDYDRDLFLGDLDAIYSLRDQAMPDEVYMPQPP
/translation="
RECEYAKRGDPLTECANFVRYLOFHNHNTLACGTGFOPTCALIVGIRGCHVH
PSGVEGGRGCPHEPSRPFASFTIDGELITGLTADLIGRAMIFRSGGRPALPD
OSLIHDPFRVMAARIPENSDDNDKRVYFEFSSEIYVSPDGSNHNIVTSGRVAND
GOVILVKNKSTFLKARIVCSYVPGGATGHEDDLEFLMPPAGKSLVVALFST
AVQGGFANCVYHADIWEFNENPFAHRGDPQHOHMGPGGVGAPPFGPGVCSMTK
PFGSTKIDYDEYLDQFARAHPLFMFVPRVPRHGRIVLAKTHLAQDLQIIVYKREAE
TYVIVFLGTDGSGVLYLVALQSGSAEEVEVLELVFPVPIPIEMELISAKQD
VGSRLVGLDRLGHQCTVTGTACAECLARDPYCAMDAGSCNHRPSLGRKRRQD
HGNDALQDCLQSQEEAEGLVAIVATVYGIENHSFPLDCSPDAVRMLLRPGG
PDQKRGDLVHTREGLFRRLSRFDAGVYCTULHEGFSQTVYRALIAYVASOL
LPEPFRKPRPAPRGGLASTPPKAWKYLIDLGHNTPRDELCEYVNCRTTEB
LPEPFRKPRPAPRGGLASTPPKAWKYLIDLGHNTPREVAAT

```

BASF COUNT	937 a	1220 c	1407 g	2023
ORIGIN				
Query Match		82.9%	Score 2177.6;	DB 9; Length 4469;
Best local Similarity	100.0%;		Pred. No. 0;	
Matches 2177; Conservative		1;	Mismatches	0; Indels 0; Gaps

[illegible][illegible]

10

TITLE Human semaphorin 2smf-16
 Patent: WO 0140278-A 1 07-JUN-2001;
 JOURNAL ZymoGenetics, Inc. (US)
 FEATURES Location/Qualifiers

source 1.2340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1.2340
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC42673.1"
 /db_xref="GI:14536656"
 /translation="MAPSAMAICWLLGILLHSGSSGSPSPYPRILRLRYGAMVRK
 PSTMMETEFRLYLSANRSAIFLGPOSLMLOMYLDEXRDLFLGDLALSLKLD
 OAMPDRTEPCANFVRLQPHNRHLACGTGALOPTICALITVGRHEVHLHEPSV
 ESGRGCPHEPRPASTFIDELTYGLTADFLGRBEMHNVSRVGVNDAGGQRY
 HDRPYMARIPENDSDNDKYFFPSETPSPDGSNHYTVSRVGVNDAGGQRY
 LVNKSSTPKARLYVSPGGAETHFDLEDVLLMKAKSLFVYALFSTYSAVQ
 GFAVCYTHADITVEVNGPFAHSDGPDQWQVPGKVPFPRPGVCSKMTAOPRPG
 STADYDDEVLYQFARAHPLMFPVPRHGRPLVKTALAOOLHOLIVDREVEDTNY
 IFLETDGSLVLYALQAGSAPSEVEVLELOVKEVPTITEMELSVKRMALYGRN
 LGVALRLHOCETVGTACAECLARDPCAMPAGACTHYRSLGRREFRDLKRN
 ALQICGSOEEAVGLVAATVYGTETHTNTEPELCKPSQAVRWMLLORPDEBPV
 KTDRELVATPTEGTLFRLSRFDAGTTCITTEFLHNPVDEYCRVWCNRTTSCGFRS
 EKPEPPPARGLASTPDKAMKRDILQLGFANLPVDEYCRVWCNRTTSCGFRS
 RSRKQARCKSMAGLELGKMKSRVHAHNRITREVEAT"

BASE COUNT 406 a 748 c 743 g 443 t
 ORIGIN

Query Match 81.8%; Score 2148.6; DB 6; Length 2340;
 Best Local Similarity 95.1%; Pred. No. 0; Indels 117; Gaps 2;
 Matches 2285; Conservative 1; Mismatches 0

280 ATGGCCCCCTGGGCTGGGCAATTTGCTGCTAGGAGGCGCTCTCTCCATGGAGGT 339
 1 ATGGCCCCCTGGGCTGGGCAATTTGCTGCTAGGAGGCGCTCTCTCCATGGAGGT 60
 340 AGCTCTGGCCCCAGCCCCGGGCGGAGTGTGCCCGGCTCTCTCTCTCTCTCTCTCT 394
 61 AGCTCTGGCCCCAGCCCCGGGCGGAGTGTGCCCGGCTCTCTCTCTCTCTCTCTCT 120
 395 -----ACCT 405
 121 ATGTCCGAAAGCTTCAGACACATGTGATGAAACATTTTCAGATACCTCTCTCTCT 180
 406 GCCAACCCCTCTGCACTCTTTCTGGGCCCCAGAGGCTCTCTCTCTCTCTCTCTCTCT 465
 181 GCCAACCCCTCTGCACTCTTTCTGGGCCCCAGAGGCTCTCTCTCTCTCTCTCTCTCT 240
 466 CTAGATAGTACGAGACCGCTCTTTCTGGGGTGGCTGAGACGCCCTCTCTCTCTCTCT 525
 241 CTATATAGTACGAGACCGCTCTTTCTGGGGTGGCTGAGACGCCCTCTCTCTCTCTCT 300
 526 CTGACACGAGGATGGCAGATCCCGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 329
 301 CTGACACGAGGATGGCAGATCCCGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 329
 586 GAGGAGTGTGTTGAAAGGAAAGATCTTTGACAGAGTGGCCCACTTGTGGGGGTG 645
 330 -----GACAGAGTGGCCCACTTGTGGGGGTG 357
 646 CTAGACGCTCACACCGGACCCACCTGTAGCTGTGGGCACTGGGGCTCTCTCTCTCTCT 705
 358 CTAGACGCTCACACCGGACCCACCTGTAGCTGTGGGCACTGGGGCTCTCTCTCTCTCT 417
 706 TGTGCCCTCATCAGATTGGCCACCGTGGGAGCATGTCTCACTCTGAGAGCCCTGGAGT 765
 418 TGTGCCCTCATCAGATTGGCCACCGTGGGAGCATGTCTCACTCTGAGAGCCCTGGAGT 477
 766 GTGGAAGTGGCCGGGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 825
 478 GTGGAAGTGGCCGGGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 537

826 ATGAGAGGAGCTGTACACGAGGCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 885
 538 ATGAGAGGAGCTGTACACGAGGCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 597
 886 TTCCGAGTGGAGGCT 945
 598 TTCCGAGTGGAGGCT 657
 946 GAGCCCGGCTTGTGATGGCCCGGAGTCTCTGAGAGTCTGACACGAGACATGACAAAG 1005
 658 GAGCCCGGCTTGTGATGGCCCGGAGTCTCTGAGAGTCTGACACGAGACATGACAAAG 717
 1006 GTGATCTTCT 1065
 718 GTGATCTTCT 777
 1066 GTGATCTTCT 837
 778 GTGATCTTCT 1185
 1126 AATGAGAGCTTCT 897
 838 AATGAGAGCTTCT 1245
 1186 GAGACCCACTTGTACACGCTAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957
 898 GAGACCCACTTGTACACGCTAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1305
 1246 CTGAGAGTGTACGCGCTGTCTACACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
 958 CTGAGAGTGTACGCGCTGTCTACACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
 1306 GTGATCCACATGAGAGACATCTGGAGAGTCTTTCACAGGCGCTTGTCCACAGAGATGG 1365
 1018 GTGATCCACATGAGAGACATCTGGAGAGTCTTTCACAGGCGCTTGTCCACAGAGATGG 1077
 1366 CCTCAGACCAATGGGGGCTCTATGGGGGCAAGGTCCTTCTCTCTCTCTCTCTCTCTCT 1425
 1078 CTGACACCAATGGGGGCTCTATGGGGGCAAGGTCCTTCTCTCTCTCTCTCTCTCTCT 1137
 1426 CCTCAGACCAATGGGGGCTCTATGGGGGCAAGGTCCTTCTCTCTCTCTCTCTCTCTCT 1485
 1138 CCTCAGACCAATGGGGGCTCTATGGGGGCAAGGTCCTTCTCTCTCTCTCTCTCTCTCT 1197
 1486 GAGTGTCTGCAATTTGCCGAGCCCAACCCCTCATGTTCTGGCTGTGGCTCTCTCTCT 1545
 1198 GAGTGTCTGCAATTTGCCGAGCCCAACCCCTCATGTTCTGGCTGTGGCTCTCTCTCT 1257
 1546 GCGCGCTCTGCTTGTCAAGACCCACCTGTGGCCACAGCTACACAGATGTGGTGGAC 1317
 1258 GCGCGCTCTGCTTGTCAAGACCCACCTGTGGCCACAGCTACACAGATGTGGTGGAC 1665
 1606 GCGGTGAGGACAGAGAGTGGGACTTACAGATGATTTTCTGGGGACTGACTCAAGGCT 1377
 1318 GCGGTGAGGACAGAGAGTGGGACTTACAGATGATTTTCTGGGGACTGACTCAAGGCT 1725
 1666 GTGCTCAAGATGATGCT 1437
 1378 GTGCTCAAGATGATGCT 1785
 1726 GAGGAGCTCAGAGTGTATGAGTGGCAACACCTATACACGAAATGAGATCTGTCAAA 1497
 1438 GAGGAGCTCAGAGTGTATGAGTGGCAACACCTATACACGAAATGAGATCTGTCAAA 1845
 1786 AGGCAATGCTATAGTGGCT 1557
 1498 AGGCAATGCTATAGTGGCT 1905
 1846 GAGACTTACAGGACATGCT 1617
 1558 GAGACTTACAGGACATGCT 1965
 1906 GATGTGCT

QY	280	ATGCCCCCTCGGCTGGCCCA	Indels	Gaps
			15;	2;

RESULT 7	AX155172		
LOCUS	AX155172		
DEFINITION	Sequence 3 from Patent WO0140278.	2337 bp	DNA
VERSION	AX155172		linear
KEYWORDS	AX155172.1 GI:14536657		PAT 22-JUN-2001
SOURCE			
ORGANISM	synthetic construct.		
REFERENCE	synthetic construct.		
AUTHORS	artificial sequence.		
TITLE	1 (bases 1 to 2337)		
JOURNAL	Holloway, J.L. and Foley, K.P.		
	Human semaphorin zsmf-16		
	Patent: WO 0140278-A 3 07-JUN-2001;		
FEATURES	Zymogenetics, Inc. (US)		
source	Location/Qualifiers		
	1..2337		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="Degenerate Polynucleotide sequence of ZSMF-16"		
BASE COUNT	308 a 273 c 468 g 293 t 995 others		

Y	2377	TCGCCACCCACCGAGCGCTGGTACAGAGCAATCCCTGACAGCTTATGGCTTGCCCACTG	2436
Db	2089	MSNACNCCNCCNARAGCTGTGATPAARGAATHTTTCACATRYTATAGNTTGYCNAAYTN	2148
Oy	2437	CCCCGGGGGATGATGATCTGTGACGGGGTGTGTGCGAGGGGCGACCCAGAACTGCACGGC	2496
Db	2149	CCNNNGTNGTNGAYGARTAYTGAGRMNGTGTGTGTGNGNACNACNARFGYNGSGN	2208
Oy	2497	TGCTTCGAGAGCCGAGCGCGGGGCGACAGCCGAGGGGCAAGAGCTGGCGAGGCTGGAG	2556
Db	2209	TGCTTGMGMSMGMWSMNGMNGMNAARCARCCNNGNNGMNAFMSNTGGCGNGNTTNGAR	2268
Oy	2557	CTAGCGAAGAGATGAAGAGCCGGGTGCATCCCGAGGACATCGAGAGCCCGCGGAGGTG	2616
Db	2269	YTGGAANAARATGATGAARWSMNGTNGTAYGCGNARCAIYAATMGNACNCCMGNGAGTIN	2328
Oy	2617	GAGGCCAC	2624
Db	2329	GARGCNC	2336
RESULT	8		
LOCUS	AF034744	2898 bp	linear
DEFINITION	Mus musculus semaphorin	2898 bp	mRNA
ACCESSION	AF034744	M-Semak	mRNA, complete cds.
VERSION	AF034744.1	GI:4104333	
KEYWORDS			
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 2898)		
JOURNAL	Miyazaki, N., Furuyama, T., Sakai, T., Fujioaka, S., Mori, T., Ohoka, Y.,		
MEDLINE	Takeda, N., Kubo, T. and Inagaki, S.		
PUBMED	Developmental localization of semaphorin H messenger RNA acting as		
AUTHORS	a collapsence factor on sensory axons in the mouse brain		
TITLE	Neuroscience 93 (1), 401-408 (1999)		
JOURNAL	2 (bases 1 to 2898)		
MEDLINE	Miyazaki, N., Furuyama, T. and Inagaki, S.		
PUBMED	Submitted (16-NOV-1997)		
AUTHORS	Sciences Osaka University Faculty of Medicine, 1-7 Yamadaoka,		
TITLE	Suita, Osaka 565, Japan		
JOURNAL	Location/Qualifiers		
MEDLINE	1..2898		
PUBMED	/organism="Mus musculus"		
AUTHORS	/strain="C57/Black 6"		
TITLE	/db_xref="taxon:10090"		
JOURNAL	370..2697		
MEDLINE	/codon_start=1		
PUBMED	/product="semaphorin H-M-Semak"		
AUTHORS	/protein_id="A010196.1"		
TITLE	/db_xref="GI:4104334"		
JOURNAL	/translation="MAPAGHLLTLLMGHLLLEMTPOGHSANPSYPRLSRSHKLELLEIN		
MEDLINE	PSISQSPGLGFLDTLLMLDEYERLTVGGRDLYTSLNERSVDGRIEYMTSTAVK		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
JOURNAL	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
MEDLINE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
PUBMED	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
AUTHORS	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		
TITLE	EECTIKMGDANCANYIVLHYNRTHLLTCATGAPPHCAATRVGHSHSEEPFLHES		

QY	274	CCGAGCATGCCCCCTTCGCGCTGGCCCATATTTGCTGCTCTAGAGGGGCTCTGCTGCTCAT	333
Db	364	CGCACCATTGACACCGGCGGACACATCTCACACCTTGTCTGCTGGGGTCACTGCTGGAA	423
QY	334	GGGGGAGCTCTGGCCCCACGCCCCCGGCACTGTGGCCCCCGGCTGCTCTACCGA	393
Db	424	CTTGGAGCCCAAGGTACACCGCGGAACCCCTCTACCCAGGCTACGCTCTCAATAAA	483
QY	394	GACCTCTGTGCGCAACGCGCTGCTGCATCTTTCTGGGCCCCCGGCTCCCTGAACCTC	453
Db	484	GAATTTTGGAACTGAATAGACCTTCAATATTTCAAAACCCCTTGGATTTCTTGAATCT	543
QY	454	CAGGCGCATGTACCTGATGATGATACCGAGACCGCTCTTCTGGGTGGCTGGAGGCCCTC	513
Db	544	CATACAAATGCTGCTGGATGATATCAAGAACGGCTCTTGTGGAGGACAGAACCTTGTCT	603
QY	514	TACTCTGCGGCTGGAGCAAGCATGGCGCAAGATCCCGGAGGCTCTGTGGCCAGCGAG	573
Db	604	TATTCCTGAACTTTGGAAAGAGTCAAGTGAAGGCTACAGAGAGATATATCTGGCGAGACA	633
QY	574	CCAGAGACAGAGAGAGAGTGTGTTCCAAAGGAGAGATCTTTGACAGAGTGGCGGCTC	633
Db	664	GCAGTAAAGGATGAAGATCATATGATGAAGAAAGAGCGC---AAATGAGTGGCAAT	720
QY	634	TTGCTGGGGGTGTACAGCTTCAACACGAGACCTGCTAGCCGTGGGCTGGGCGCC	693
Db	721	TATATCGGGTTTGTGATCTACTACAAACAGACACACTTCTGACCTGTGCTAGTGA	780
QY	694	TTCCAGCCCACTGTGCTCCCTCATACAGTGGCGCACCGT---GGGAGCATGTGCTCCAC	750
Db	781	TTTGATCTCAACTGTGCTCTTCAAGAGTGGCGCACCTTCAAGAGAACCCCTTTTCCAC	840
QY	751	CTGGAGCCCTGGAGTGTGGAAGTGGCGGGGGGCGGTGCTCTCAAGAGCCAGCGCTCC	810
Db	841	CTGGAGTCTACACAGATCTGAGAGAGAGAGAGGCGCAAGATGCTTTTGAACCCCACTCTCC	900
QY	811	TTTGGCAGCACTTCAATGAGCGGGAGCTGTACACGGGTCTCACTGTGCTGCTCTGGG	870
Db	901	TTTGTGCTCAGCTGATGTTGGGATGAGCTGTGTTGCTGACCTTACACTGACTATTTGGGG	960
QY	871	CGAAGGCCATGATCTTCCGAAGTGGAGTCTCTGGCCACTCTGCGTTTCCGACTCTGAC	930
Db	961	AGAGACTGTGGGATCTTCCGAGCATGGGGAATGAGCCATATTTGCACTGAGCATGAC	1020
QY	931	---CAGATCTCTTGCACGACCGCCCGTGTGTATGATGCGCCGATTCCTAGAACTCT	987
Db	1021	GATGAGCGGCTCTGAAAGAACCAAAATTTGTAGTTCAATATATGATTCCTGATTAACGA	1080
QY	988	GACCAGAGCAATACAAAGGTATCTTCTTCTCTGAGAGCGTCCCGTCCGCGCATGCT	1047
Db	1081	GACCGAGATGACACAAATGTATCTTTCTTACTGAGAAAGCGCTCGAGCGGAGAAC	1140
QY	1048	GGGTCAACCATATGCTACTGTGACCGCGTGGGCGCGCTCTGCTGATGATGCTGGGGGC	1107
Db	1141	AAGCGCCACACGATCTA---CACCGAGTGGGGCGCGCTGTGCTGTAAGCATGGAGGA	1197
QY	1108	CAGCGGTGCTGTTGAACAAATGAGCACTTCTCAAGGCGAGCTGTGCTGTGCTGGT	1167
Db	1198	CAGAGAAATCTGTGTAACAAAGTGGACATTTCTTAAAGCGCGGTGGTTGTGCTGAGT	1257
QY	1168	CGCGGCGCTGTGTGCGGAGACCACTTTGACAGCAGTAAGAGATGTGCTGCTGTGG	1227
Db	1258	CGGGATGATGATGATGATGACACATCTTTGACAGCATAGAGATGTGTTTACTGCGG	1317
QY	1228	CCCAAGGCGGAGAGAGCTGAGAGTGTAGCGCTGTTCAGCAACCGTCACTGCGGTCTC	1287
Db	1318	ACCGAGATCTTAAAGTATCCAGTATTTTGTGACTGTTTAATCTACACGACATATATTT	1377
QY	1288	CAGGCTTCCCGCTGTGTGTCACATGCGACACATCTGGAGGTTTTCAGAGGGCC	1347

Db 2449 ATGCGTCCCTCCCTTAGCGGTATGTCTCAGGGGACAAACCGTGTACAGAAATTC 2508

QY 2410 CTCAGCTCATTTGGCTTGGCCCAACCCCGGGGTGATGATGAGCGGCTGTGGT 2499

Db 2509 TTGCAGCTGATTTGGCTTACAGCACTTCCAGAGAGTGGAAATATCTGCGAAAGGTGG 2568

QY 2470 TGCA 2473

Db 2569 TGTG 2572

RESULT 9

LOCUS MUSEMH 2328 bp mRNA linear ROD 20-MAY-1998

DEFINITION M. musculus mRNA for semaphorin H.

ACCESSION 280941

VERSION 280941.1 GI:3150195

KEYWORDS semaphorin H.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 2328)

AUTHORS Christensen C.R., Klingelhofer J., Tarabiykina S., Hulgard E.F., Kramrov D. and Iukandina E.

TITLE Transcription of a novel mouse semaphorin gene, M-semah, correlates with the metastatic ability of mouse tumor cell lines

JOURNAL Cancer Res. 58 (6), 1238-1244 (1998)

98175564

2 (bases 1 to 2328)

Christensen, C.C.

Direct Submission

Submitted (11-Oct-1996) Christensen C. C., Danish Cancer Society.

Molecular Cancer Biology, Strandboulevarden 49, Copenhagen, Denmark, 2100

revised by [3]

3 (bases 1 to 2328)

Christensen, C.C.

Direct Submission

Submitted (14-Apr-1998) Christensen C. C., Danish Cancer Society.

Molecular Cancer Biology, Strandboulevarden 49, Copenhagen, Denmark, 2100

On May 22, 1998 this sequence version replaced gi:1619599.

Location/Qualifiers

1..2328

/organism="Mus musculus"

/strain="Balb/c FC3H"

/db_xref="taxon:10090"

/cell_line="66c14"

1..2328

/codon_start=1

/product="semaphorin H"

/protein_id="CAB02590.1"

/db_xref="GI:3150196"

/db_xref="SPTREMBL:P0275"

/translation="MAPAGHLITLLWGHLLLEMTPOHSANPSYPRLSHKELEFLN
RSTQSPGLGLDHTMLDDEYERLIVGGRDLYVSLNLRVSDGRLVSTWSTAYKV
ECTKMGKADNACANVIRVHLHYNRTLLICATGAPDFCAFVGHSHSEPLEHED
HSEKRGKCPDPNRSFVSTLVNGLFAGISYWGSDAIFRSMKLOHITREHD
BLIKERFVGSIMLPDEBDNKNMFEPEKLEENNAHTLYTVGRLAYNDGG
ORILVNWSTELKARLVCSVGANGIDITVPELEDDVLLPDRPKNVIGLNTSN
IFRGHACVYHNSIREAFNGPIVAKRBPETVHSLYSGKVPKPGSCAKVNGKYG
TKGDPAIDIRFARIDPLMOPIKRVHKKPLVTDGKYNIROLANRVAEDGOYV
LFGDGTGLKLVITYNOETEMMEVLELELQIFKDPAPLISHSKROLYTISA
SAVORRHCHMGYSACADCLADPYCAMDGISCSRYPLTGAHARRRRRODVRG
NAAOCFGQDFGADLDRPEERLAGEISNTLECPKPSLOAVLWQKRGVKE
EYKIDRVYKMLDGLITRVKRSADGTECCOVEHNFTVYRKITLLEVEFHVKEFM
HKDHEERHKKPCPLSLMSGCTPKPKFELIOLIGYENFORVEYCEKWCIDRKKR
KLMSPSKRVANPDEKRLRSKAEHPRRLHLLS"

BASE COUNT 665 a 542 c 595 g 526 t

Query Match 24.0%; Score 629.6; DB 10; Length 2328;

Best Local Similarity 57.6%; Pred. No. 8.4e-107; Matches 1278; Conservative 1; Mismatches 900; Indels 39; Gaps 7;

QY 280 ATGCGCCCTCCGCGCTGGGCCATTTGCTGTGCTGCTAGAGGGCCCTCCTCATGGGGGT 339

Db 1 ATGCGACCGCGCGGACATCTCCTCACTTGTGCTCTGCTGGGGCTACTGCTGGAACTCG 60

QY 340 AGCTTGGCCCGCCAGCCCGCCCAAGTGTGCGCCCTGCGGCTCTCTACAGACCTC 399

Db 61 ACCCGAGTCTACTCCGCGAACCCTCTCTACCCAGGCTACCGCTGCACATTAAGAACTT 120

QY 400 CTGCTGCCAAGCGCTCTGCGCATCTTCTGCGGCCCGCCAGGCTCTCCCTCAACCTCAGGCC 459

Db 121 TTGAACTGATAGACTTCATATTTCAAGCCCTTGTGCGGCTCTGATCTTATCATACATA 180

QY 460 ATGTACTAGATGATGATCGGACCGGCTCTTCTGCTGGCGCTTGACGCGCTCTACTCT 519

Db 181 ATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

QY 520 CTGCGGCTGACACGAGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 579

Db 241 CTGAACCTTGGAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 580 CAGAGGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639

Db 301 AAGGTAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357

QY 640 CGGAGTCTACACCTTCACACACCGGACCGGACCTGCTAGCTGTGCTGCTGCTGCTGCTGCT 699

Db 358 CGGCTTTGATCTACATACACACGAGACACACCTTGTGATGATGATGATGATGATGATGAT 417

QY 700 CCCACTGTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756

Db 418 CCACACTGTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477

QY 757 CTTGCGAGTGTGAG 816

Db 478 TCACACAGATCTGAG 537

QY 817 AGCATCTTCAATACAGCGGAG 876

Db 538 TCCACGCTATGAG 597

QY 877 GCCATGATCTTCCGAG 933

Db 598 TCGCGCATCTTCCGAG 657

QY 934 AGTCTTTCAG 993

Db 658 CGGCTCTGAG 717

QY 994 GACATGACAG 1053

Db 718 GATGACACAAAT 777

QY 1054 AACCATGATCTGTAG 1113

Db 778 CACACATCTA---CACCGAGTGGGCGGCTGTGCGGATGATGATGATGATGATGATGATGAT 834

QY 1114 GTGCTGTGAG 1173

Db 835 ATCTGTGAG 894

QY 1174 CTTGCTGTGAG 1233

Db 895 ATGAATGAGATGAG 954

QY 1234 GCGGAG 1293

Db 955 GATCTGAG 1014

QY 1294 TTGCGCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1353

Db	1376	CCGAGCAGAA	-----CCTTTGGACCTTTCAGTTCACCAAGACCTTCCAGAC	1423
Db	1376	CCGAGCAGAA	-----CCTTTGGACCTTTCAGTTCACCAAGACCTTCCAGAC	1423
QY	1486	GAGGTGTGCAATTTGGCCGAGCCACCCTCATATGTTCTGGCCTGTGGCCCTGCACAT	1545	
Db	1424	GATGTATTCACCAATTTGGCCGAGACCACTCCCTCATATGTTCTGGCCTGTGGCCCTGCACAT	1483	
QY	1546	GGCCGCGCTGTGCTTGTCAAGAACCACTGGCCGAGCAGCTACACAGATGTGGTGGAC	1605	
Db	1484	GGGCGCCCTCTTTCTCTACAGTTGAGGCCAATTCACCTTACTCAATTCACCTGGCGGAC	1543	
QY	1606	CGGTGTGAGGAGCAGAGATGGGAGCTACGATGATATTTCTGGGAGCTGACAGGCTCT	1665	
Db	1544	CGGTGTGAGGAGCAGAGATGGGAGCTACGATGATATTTCTGGGAGCTGACAGGCTCT	1603	
QY	1666	GTGCTCAAAATCATATGCTCTCCAGGAGGGGCTCAGCTGACATGAGGAGAGGTTCTG	1725	
Db	1604	GTGCTCAAAATCATATGCTCTCCAGGAGGGGCTCAGCTGACATGAGGAGAGGTTCTG	1663	
QY	1726	GAGGAGCTCCAGTGTTTAAGGTGGCCAAACATATACCAAGGAGATCTGTGCMAA	1785	
Db	1664	GAGGAGCTCCAGTGTTTAAGGTGGCCAAACATATACCAAGGAGATCTGTGTCCAA	1723	
QY	1786	AGGCAATATGATATAGTGTGGCTCTCGGCTGGGTGGCCCAAGCTGGGGGTGCACATGT	1845	
Db	1724	AGGCAATATGATATAGTGTGGCTCTCGGCTGGGTGGCCCAAGCTGGGGGTGCACATGT	1783	
QY	1846	GAGACTTACGCACTGCTGTGTCAGAGTGTGCTGGCCGCGGAGCCCACTATGTCCTG	1905	
Db	1784	GCTGCGCCACGCGCGGCTGTGACAGCAATGCTGTGTGGCGCTGCACCTTACGCGCTGG	1843	
QY	1906	GATGTGTGCTCTCTTATACCACTACCGCCCAAGCTTGGCAAGCGCGGTTTCCGCGGAG	1965	
Db	1844	GATGTGTGCTCTCTTATACCACTACCGCCCAAGCTTGGCAAGCGCGGTTTCCGCGGAG	1900	
QY	1966	GACATCCGCGACGCAACCTTGCCTGCAGTGTCTGGGCGCAGAGCCAGAGAGAGCA	2025	
Db	1901	GACATCCGCGACGCAACCTTGCCTGCAGTGTCTGGGCGCAGAGCCAGAGAGAGCA	1960	
QY	2026	GTGGGACTTGTGCGACGACCATGTGCTACAGCGACGAGCACAATATGACCTTCCGCGAG	2085	
Db	1961	GTGGGACTTGTGCGACGACCATGTGCTACAGCGACGAGCACAATATGACCTTCCGCGAG	2008	
QY	2086	TGCGTCCCAATGCTCCCAAGCTGTGCTGTGCTGCTGCTTGTGCAAGGCCAGGGGATGAG	2145	
Db	2009	TGCGTCCCAATGCTCCCAAGCTGTGCTGTGCTGCTGCTTGTGCAAGGCCAGGGGATGAG	2068	
QY	2146	GGGCGCTGACGAGTGAAGAGGAGCAGAGCTTGTGACACAGGAGGGGGCTGTCTTC	2205	
Db	2069	GGGCGCTGACGAGTGAAGAGGAGCAGAGCTTGTGACACAGGAGGGGGCTGTCTTC	2128	
QY	2206	CGCAGGCTTACCGCTTTCGATGCGCGGACACCTGACACCACTTGTGAGCATGCTTC	2265	
Db	2129	CGCAGGCTTACCGCTTTCGATGCGCGGACACCTGACACCACTTGTGAGCATGCTTC	2188	
QY	2266	TGCCAGACTGTGTGCTGCTGCTGTGCTGATTTGGGCTCACAACCTGAGACACTG	2325	
Db	2189	TGCCAGACTGTGTGCTGCTGCTGTGCTGATTTGGGCTCACAACCTGAGACACTG	2248	
QY	2326	TTCCTTCGCGAG	2337	
Db	2249	TTCCTTCGCGAG	2260	
RESULT 11				
BC013975				
LOCUS	BC013975	2825 bp	mRNA	linear
DEFINITION	BC013975	Homo sapiens, Similar to sema domain, immunoglobulin domain (I9), short basic domain, secreted, (semaphorin) 3b, clone MGC:12697		
ACCESSION	BC013975	IMAGE:4121913, mRNA, complete cds.		
VERSION	BC013975.1	GI:1559240		
KEYWORDS	MGC.			
SOURCE	human.			

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 2825)

JOURNAL

Strausberg, R.

Direct Submission

Submitted (10-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT

Contact: MGC help desk

Tissue

Email: cgapbs-remail.nih.gov

CDNA

Tissue Procurement: ATCC

DNA

CDNA Library Preparation: Rubin Laboratory

Sequencing

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Sequencing

Sequencing Center (NISC), National Institutes of Health Intramural

Web site:

<http://www.nisc.nih.gov/>

Contact:

nisc_mgc@hgti.nih.gov

Shevchenko, Y.

Wetherby, K.D., Beckstrom-Sternberg, S.M.,

Benjamin, B.

Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,

Dietrich, N.T.

Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,

Lim, M.

Maduro, Q.L., Masello, C., Mastrian, S.D., McCluskey, J.C.,

McDowell, J.

Pearson, R., Snyder, B., Stantilpop, S., Thomas, P.J.,

Tiongson, E.E.

Touchman, J.W., Tsurgoun, C., Vogt, J.T., Walker, M.A.,

Zhang, L.

H. and Green, E.D.

FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

location

Series: IRAL Plate: 17 Row: f Column: 20

passed

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4759091.

1. 2825

Location/Qualifiers

/organism

/db_xref="taxon:9606"

/clone

/feature="MGC:12697 IMAGE:4121913"

/feature

/feature="Muscle, rhabdomyosarcoma"

/clone

/clone="NIH-MGC_17"

/lab

/lab="host=DH10B-R"

/note

/note="Vector: pOTB7"

82. 2328

/product="similar to sema domain, immunoglobulin domain (Ig) short basic domain, secreted, (semaphorin) 3B"

/protein

/protein="GI:15559241"

/translation

/translation="MGRAGAAAVIPGLALLMAVGLGSAAPSPRLRISFQDLQAHGLQTFLEPTCCYQALVLEDERGLFVGAENHVASLNLNISRKRLAMPVAVEMREBCNMAGDITGECMNFVILHAYNRTHLLACGATHTPCAFVGVHRAEPEVLRADPERIEDGKSPDPRHRAASVYGEELVSGVADLDLGRPTTFRSAGSPSLRTEHDSRLVNEPKFVYVPIPESENPPDDKTYFFRETVAAAPALRSLSVRGOICRNDVGOCSANCYVSMNDVRAFLGPRFAKDEGPMHMOVSGRVYRPPGCKPSKTEGSTRDPTDVTIVKIVSVKSPKSAEGLLLELHVEEDSAATSMOISKRQOLYVARSVAQIALHRCAGARVCTECCILADPCAMDVACVGRPSAKRRRRDVRNDRVSLTSSGDSRPALHLEHKEVGVSSAPLECEPRISQAVMTFQAGTATHTOYVLEERTRTAGRLRLRLRRDSGYVILCAVROGTPQPLRSLHVLSTQAEIRLAREEAPAPGPKIWRDFQLVLEPVGSGSANSURKRCRPOPALSLPLESRGRNRRTTHAPRAERGPSATTHM"

BASE COUNT

574 a 867 c 924 g 460 t

ORIGIN

Query Match

22.7%; Score 595.6; DB 9; Length 2825;

Best Local Similarity

59.2%; Pred. No. 1.6e-100;

Matches 1179; Conservative

1; Mismatches 770; Indels 42; Gaps 8;

QY

352 AGCCCCGCCCCAGTGTGCCCCCGCTGCTCTACCGAGAGACTCTGTCTGCCAAC 411

Db

148 AGTGGCGCCCCCGCCCCCGCTTGGCTCTCTCCACAGAGCTCCAG--GCCTGG 204

QY 412 CGCTTGCCATCTTCTTGAGGCCGCCGAGGCTCCCGAAGCTTCAGAGCATGTACTAGAT 471

Db 205 CATGCTCTCCAGACTTCTTACGCTGAGGAGCACTCTCTCTACAGGCGCTTGCTGGAT 264

QY 472 GAGTACGAGAGGCGGCTCTTCTTGCTGCTGCTGAGGCGCTCTCTCTCTCTCTCTGAG 531

Db 265 GAGAGGCTGAGGCGGCTCTTCTTGCTGCTGCTGAGGCGCTCTCTCTCTCTCTCTGAG 324

QY 532 CAGCAGTGGCCAGATCCCGGAGGCTCTGAGGCGGAGGCGGAGGCGGAGGAGGAG 591

Db 325 AACATCAGCAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 384

QY 592 TGTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 651

Db 385 TGCAACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 444

QY 652 CCTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711

Db 445 GCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504

QY 712 CTGATCAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 768

Db 505 TTGTGTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564

QY 769 GAAAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 828

Db 565 GAGATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624

QY 829 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 888

Db 625 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684

QY 889 CGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 945

Db 685 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 744

QY 946 GACCCCGGTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005

Db 745 GAGCCAGGTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 804

QY 1006 GTGTCT 1065

Db 805 ATCTACT 864

QY 1066 GTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125

Db 865 GTGTCCGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924

QY 1126 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981

Db 925 AAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981

QY 982 GAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041

Db 1042 CCGCTGCT 1098

QY 1306 GTGTACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1365

Db 1099 GTGTACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158

QY 1366 CCGACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1425

Db 1159 CCGATGACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218

QY 1426 CCGACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1485

Db 1219 CCGACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266

us-09-813-290-1.rge

Page 17

[illegible][illegible]

Db 600 GGGGAGAGATATCTAGGGGTGGCAGACCTCATGGAGAGACTTATACATCTTT 659
 QY 889 CGAAGTGGAGTCTCGCCAGAGCTCTGCTTCCGACTGACACAGATCTTTCAC 659
 Db 660 CGCAGCTAGAGGACAGCTCCAAAGTCTCCAGACAGACAGATCTCCGCTGCTCAAT 719
 QY 946 GACCCCGGTTTGTATGAGCCCGGATCCCTGAGAACTCTGACAGCAAGCAATGACAG 1005
 Db 720 GACCCCAAGTGTATGAGTATTTGATCCGAGAGCCGAGAACCCAGACAGCAAGCA 779
 QY 1006 GTGTACTTCTTCTCGAGAGAGTCCCTCGCCGATGATGATGATGATGATGATGAT 1065
 Db 780 ATCTACTTCTTCTCGAGAGAGTCCCTCGCCGATGATGATGATGATGATGATGAT 1065
 QY 1066 GTACAGCCGCTGGGCGGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 839
 Db 840 GTGTCGCGGCTGGGCGGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1125
 QY 1126 AATGAGAGCACTTCTCGAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGAT 899
 Db 900 AATGAGAGCACTTCTCGAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGAT 1185
 QY 1186 GAGACCCACTTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 956
 Db 957 GACACCCACTTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1245
 QY 1246 CTGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1016
 Db 1017 CCGTGTCTTATGCTGCTTCTTCCAGAGTCTGCTGCTGATGATGATGATGATGATGAT 1305
 QY 1306 GTGTAGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1076
 Db 1077 GTGTAGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1365
 QY 1366 CTGTAGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1136
 Db 1137 CCATGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1425
 QY 1426 CCGAGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1195
 Db 1196 CCGAGAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1485
 QY 1486 GAGGTGTAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1243
 Db 1244 GATGTAGAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1545
 QY 1546 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1303
 Db 1304 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1605
 QY 1606 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1363
 Db 1364 CCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1665
 QY 1666 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1423
 Db 1424 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1725
 QY 1726 GAGGAGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1483
 Db 1484 GAGGAGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1785
 QY 1786 AGGCAATGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1543
 Db 1544 AGGCAATGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1845
 QY 1846 GAGCAATGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1603
 Db 1604 GCGGCGCTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1905
 QY 1906 GAGGTGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1663
 Db 1906 GAGGTGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1965

Db 1664 GACGGGTGCGGTGACGCGCTTCCAGCCAG---TGCCAGAGGGGCTTCCGGGCGCA 1720
 QY 1966 GACATCGGCGCAAGGCAACCCCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1720
 Db 1721 GATCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
 QY 2026 GTGGAGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1780
 Db 1781 CTGGAGC-----ACAGGTGTGTGCGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1828
 QY 2086 TGCTGCGCCAGATGCTTCCAGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1828
 Db 1829 TGAGAGCTGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2145
 QY 2146 GGGCTGACAGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 1888
 Db 1889 GCGCAGACCCAGAGTGTGTAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGAT 2205
 QY 2206 CGCAGCTTACCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1948
 Db 1949 CGCAGCTTACCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2265
 QY 2266 TCCAGAGTGTGTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2008
 Db 2009 ACAGAGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2325
 QY 2326 TTCCCTCCGAG 2337
 Db 2069 GCGCGGGCCGAG 2080

RESULT 13
 MM293948
 LOCUS
 DEFINITION M.musculus mRNA for semaphorin Hv (3988 bp). linear ROD 17-JUL-1998
 ACCESSION 293948
 VERSION 293948.1 GI:1934966
 KEYWORDS semaphorin H.
 SOURCE house mouse.
 ORGANISM Mus musculus
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Christensen, C.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1997) Christensen C., Danish Cancer Society,
 Molecular Cancer Biology, Strandboulevarden 49, Copenhagen,
 2 (bases 1 to 3988)
 REFERENCES
 1. Kramers, D., and Lukanidin, E.,
 Transcription of a novel mouse semaphorin gene, M-semah, correlates
 with the metastatic ability of mouse tumor cell lines
 Cancer Res. 58 (6), 1238-1244 (1998)
 98175564
 FEATURES
 source
 Location/Qualifiers
 1..3988
 /organism="Mus musculus"
 /strain="Balb/c/c3H"
 /db_xref="taxon:10090"
 /cell_line="66c14"
 610..2943
 /gene="msemah"
 610..2943
 /gene="msemah"
 /note="msemah"
 /note="predicted protein. The coding sequence of the
 msemah transcript is identical to the submission with Acc.
 number 280941."
 /product="semaphorin H"
 /protein_id="CAB07988.1"
 /db_xref="GI:1934967"
 /db_xref="SPTREMBL:P70275"

Thu Oct 10 09:34:59 2002

us-09-813-290-1.rge

Page 23

Qy	2438	CCCCGGTGA	TGAGTACTGTGAGCGCGTGTGTC	2473
Db	2201	AGAGAGTGAGGAGTACTGTGAAAAAGTCTGTGTA		2236

Search completed: October 9, 2002, 12:40:47
Job time : 3291 secs

QY 541 CCAGATCCCCGGGAGGTCTGTGGCCACCGAGCCAGGACAGAGGGAGGAGTGTTCGA 600

Db 541 CCAGATCCCCGGGAGGTCTGTGGCCACCGAGCCAGGACAGAGGGAGGAGTGTTCGA 600

Db
601 AAGGAGAGATCCTTTGACAGAGTGGCCCACTTCGTGCGGGTGCTACAGGCTTCAAC 660

QY 661 CGACCCACCTGCTAGCCGTGTGGACTGGGGCCCTTCAGGCCACCTGGCCCTATCACA 720
|||||
|||||

QY	721	GTTTGGCCACCGCTGGGGGAGCATGTGCTCCACACTGGAGCGCGAGAGTGGGAAATGTCGCGG	780
Db	721	GTTTGGGCACCGCTGGGGGAGCATGTGCTCCACACTGGAGCGCGAGAGTGGGAAATGTCGCGG	780
QY	781	GGGCGGTCCCTCCACGAGGCCAGCCGCTCCCTTTGCGACACCTTTCATAGAGGGGAGCGTG	840
Db	781	GGGCGGTCCCTCCACGAGGCCAGCCGCTCCCTTTGCGACACCTTTCATAGAGGGGAGCGTG	840
QY	841	TACACGGGTCCTCACTGCTGACTTCTGGGGGCGAGAGGCCATGATCTTCCGAAATGGAGGT	900
Db	841	TACACGGGTCCTCACTGCTGACTTCTGGGGGCGAGAGGCCATGATCTTCCGAAATGGAGGT	900
QY	901	CCCTGGGCGACCTCTGCGTTTCGACTCTACCAAGAGTCTTTGACACGACCCCGGTTTGTG	960
Ddb	901	CCCTGGGCGACCTCTGCGTTTCGACTCTACCAAGAGTCTTTGACACGACCCCGGTTTGTG	960
QY	961	ATGGCCGCCCGCGATCCCTCGATGAGAACTCTGACGAGGAAATATATCAACTGTAACGGGTTTGG	960

[illegible]

1021 TCGGAGACGGTCCCTCGCCCGATGGTGGTCGAACCATGTCACTCTCAGCCGGCTGGC 1080
1020 TCGGAGACGGTCCCTCGCCCGATGGTGGTCGAACCATGTCACTCTCAGCCGGCTGGC 1080

1021 TCGAGACGGTCCCTTCGCCGATGGTGGCTGAACCATGTACTGTCAAGCCGGTGGGC 1080

1081 CGCGCTGCGTGAATGATGCTGGGGGCCAGCGGTGCTGTGAACAATGGAGCACTTC 1140

1061 CCGGCTGGCGAATGATGTCGTGGGGGCCAGCGGGTCTGGTGAAACAATGGAGCACTTTC 1140

1141 CTCAAGGCCAGGGCTGGTCTCCACC 1140

b

1141 CTCAAGCCAGGCTGCTCTCGGTGCCCCGCTGTGTGTGGCAGACCCACTTTGAC 1200

[illegible]

1201 CAGCTAGAGGATGTGTCTGCTGTGCGCCCAAGCGCGGAGAGCCTCGAGCTGTACGCG 1260

1261 CTGTTGACGACCCGTCAGTGGCCGTGTTCCAGGGCTTGCCGCTGTGTGTACACATGGCA 1320

1321 GACATCTGGAGGTTTCAACGGGGCCCTTGGCCACGACGATCCCGGCTGCTGTGTGTACCATATGGCA 1320

1321 GACATCTGGAGGTTTCAACGGGCCCTTGCCACCGAGATGGGCTCAGCACCAGTTGG 1380

1381 GGGCCCTATGGGGCAGGTGCCCTTCCCTCGCCCTGGCGTGTGCCCGACAGATGACC 1440

1381 GGGCCCTATGGGGCAAGGTGCCCTTCCTCGCCCTGGCGTGTGCCCCAGCAAGATGACC 1440

1441 GCACAGCCAGACGGCCCTTTGGCAGCACCACGAGACTACCCAGATGAGGTGCTGCAGTTT 1500

[illegible]

1501 GCCGAGCCACCCCTCATGTTCTGGCCCTGCGGCTCGACATGGCCGCTGTCTT 1560

1561 GTCAAGACCCACCTGGCCAGACCTACACCAGATCGTGTGGACCGCGTGGAGCCAGAG 1620

[illegible]

cc The sequen

1042 GATGGTGGCTCGAACCATGTCTACTGTCAAGCCGGGTGGGCCGGCTCTGGCGTGAATGATCC

PT		/note= "secretory signal sequence"
FT	mat-peptide	67..2340
FT		/*tag= C
XX		
XX		
PN	WO200140278-A2.	
XX		
PD	07-JUN-2001.	
XX		
PF	06-DEC-2000: 2000WO-US3116.	
XX		
PR	06-DEC-1999: 99US-045560.	
XX		
PA	(ZYMO) ZYMOGENETICS INC.	
XX		
PI	Holloway JL, Foley KP;	
XX		
DR	WPI: 2001-374784/39.	
DR	P-PSDB: AAB84219.	
XX		
PT	Novel human semaphorin polypeptide, ZSMF-16, useful for treating	
PT	peripheral neuropathies Alzheimer's and Huntington's disease and	
PT	polynucleotide encoding ZSMF-16 useful for detecting genetic	
XX	abnormality and cancer	
PS	Claim 2; Page 116-121; 124pp; English.	
XX		
CC	The present sequence encodes a human semaphorin polypeptide, designated	
CC	ZSMF-16. ZSMF-16 is a neurite growth and development modulator. It also	
CC	enhances spinal cord and sensory neurite outgrowth and patterning, and	
CC	is involved in the activation and regulation of T lymphocyte suppressor	
CC	ZSMF-16 is useful for detecting a genetic abnormality or cancer. ZSMF-16	
CC	polynucleotide probes can be used to detect 3p21 loss, trisomy,	
CC	duplication or translocation associated with mammary tumour tissue,	
CC	breast tumour, liver, small intestine, bone cancers, etc.. ZSMF-16	
CC	can be used to modulate neurite growth and development and demarcate nervous	
CC	system structures. ZSMF-16 are also useful for regenerating and directing	
CC	neurite outgrowth following strokes, brain damage caused by head	
CC	injuries, paralysis caused by spinal injuries, and for treating	
CC	neurodegenerative diseases such as amyotrophic lateral sclerosis,	
CC	Alzheimer's disease, Huntington's disease, Parkinson's disease and	
CC	peripheral neuropathies, or demyelinating diseases e.g., multiple	
CC	sclerosis. ZSMF-16 also acts as a mediator of immunosuppression	
CC	and thus useful for immunosuppression.	

Sequence 3340 nt. 100%
 CC of antigen in humoral and multiple
 CC in graft and organ transplants.
 XX for immunosuppression

[illegible]

Thu Oct 10 09:35:00 2002

us-09-813-290-1.rng

526 CTGACACGAGCATGGCCAGATCCCGGAGGTCTGTGGCCACCGCAGCCAGAGAGG 585
301 CTGACACGAGCATGGCCAGATCCCGGAGG----- 329
586 GAGAGTGTGTGGAAGGAGATCTTTGACAGATGCGCCAACTTCTGCGGGTG 357
330 -----GACAGATGCGCCAACTTCTGCGGGTG 705
646 CTACAGGCTCAACACGAGCCACCTGTAAGCTGTGACAGTGGGGCTTCCAGCCACC 705
358 CTACAGGCTCAACACGAGCCACCTGTAAGCTGTGACAGTGGGGCTTCCAGCCACC 417
706 TGTGCTCTATCAAGTTGGCCACCGTGGGAGCATGTGCTCCACTGTGAGGCTGAGT 765
418 TGTGCTCTATCAAGTTGGCCACCGTGGGAGCATGTGCTCCACTGTGAGGCTGAGT 477
766 GTGGAAGTGGCCGGGGGCGGTGCTTCCAGAGCCAGCCGCTTCCAGAGCTTTC 825
478 GTGGAAGTGGCCGGGGGCGGTGCTTCCAGAGCCAGCCGCTTCCAGAGCTTTC 537
826 ATGAGAGGAGGAGCTGTACAGGGGTCTCACTGCTACTTCTGGGGGAGAGGCTATTC 885
538 ATGAGAGGAGGAGCTGTACAGGGGTCTCACTGCTACTTCTGGGGGAGAGGCTATTC 594
886 TTCCGAAGTGGAGGTCTTCCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
598 TTCCGAAGTGGAGGTCTTCCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
946 GACCCCGGCTTGTGATGAGCCGCGCGATCCCTGAGAGCTGTGAGAGCAATGACAG 1065
658 GACCCCGGCTTGTGATGAGCCGCGCGATCCCTGAGAGCTGTGAGAGCAATGACAG 717
1006 GTGTACTTCTTCTCTCGAGAGCGTCCCTGCGCCAGTGTGCTGAGCAATGAGT 1065
718 GTGTACTTCTTCTCTCGAGAGCGTCCCTGCGCCAGTGTGCTGAGCAATGAGT 1125
1066 GTGAGCGGAGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
778 GTGAGCGGAGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
1126 AAATGAGAGCTTCTCTCAAGGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
838 AAATGAGAGCTTCTCTCAAGGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
1186 GAGACCCACTTGTGACAGCAATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
898 GAGACCCACTTGTGACAGCAATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1305
1246 GTGAGAGTGTGAGCGCTGTGACAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
958 GTGAGAGTGTGAGCGCTGTGACAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365
1306 GTGTACCAATGAGCAATCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
1018 GTGTACCAATGAGCAATCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1425
1366 CTTTACAGCAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1425
1078 CTTTACAGCAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1137
1426 CTTTACAGCAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1485
1138 CTTTACAGCAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1545
1486 GAGGTGTGAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1257
1198 GAGGTGTGAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1605
1546 GAGGTGTGAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1317
1258 GAGGTGTGAGTGGGGGCTTATGGGGGCAAGTGGGCTTCCAGCCGAGATGGG 1317

1606 CCGTGTGAGGAGGAGATGGAGCTTACATGTCTATTTCTGAGGAGTGTGAGGAGTCT 1665
1318 CCGTGTGAGGAGGAGATGGAGCTTACATGTCTATTTCTGAGGAGTGTGAGGAGTCT 1377
1666 GTGCTCAAAATCATGCTGTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1725
1378 GTGCTCAAAATCATGCTGTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437
1726 GAGGAGCTCCAGGTGTAAAGTGTGCAACACCTATACCGAAATGAGATGTCTGTCAA 1497
1438 GAGGAGCTCCAGGTGTAAAGTGTGCAACACCTATACCGAAATGAGATGTCTGTCAA 1497
1786 AGGCAATGTCTTACAGTGGGCTCTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1845
1498 AGGCAATGTCTTACAGTGGGCTCTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1557
1846 GAGACTTACAGGAGT 1617
1558 GAGACTTACAGGAGT 1965
1906 GATGT 1677
1618 GATGT 2025
1966 GATGT 1737
1678 GATGT 2085
2026 GTGGGAGT 1797
1738 GTGGGAGT 2145
2086 GTGGGAGT 1857
1798 GTGGGAGT 2205
2146 GTGGGAGT 1917
1858 GTGGGAGT 2265
2206 GTGGGAGT 1977
1918 GTGGGAGT 2325
2266 GTGGGAGT 2037
1978 GTGGGAGT 2385
2326 GTGGGAGT 2097
2038 GTGGGAGT 2445
2386 GTGGGAGT 2505
2098 GTGGGAGT 2557
2446 GTGGGAGT 2217
2158 GTGGGAGT 2277
2506 GTGGGAGT 2625
2218 GTGGGAGT 2337
2566 GTGGGAGT 2337
2278 GTGGGAGT 2337
2626 GTGGGAGT 2337
2338 GTGGGAGT 2337

Db 1678 TATGCTATAAAGAGCCCTGATACACGTCCTATATATGAAGAAAGTCCCTAC 1737
 QY 1408 CCTGCGCCCTGGCGTGTGCCCGACGACATATGACCGACAGCGAGCGCTTTTGGAGC 1467
 Db 1738 CCAAGGCGCTGTCTCTGTGCACGAAAGTAAC-----GGAGGCAAGTATGGAAAC 1288
 QY 1468 ACCAAGGACATACCAAGATGAGTGTGCTGAGTTTGGCCGAGCCGACCCCTCATGTCTGG 1527
 Db 1789 ACCAAGATATACCCCGATGACGCGATCCGTTCCGAAAGATGATCTTAAATGATACAG 1848
 QY 1528 CTTGTGCGGCTTCGATGAGCGCGCTGCTGCTGTCAGAGCCACCTGGCCGACAGCTA 1587
 Db 1849 CCCATTAACCTGTTCATTAACCAATATGTAACCAAGATGAAATACAACTG 1908
 QY 1588 CACGAGATGCTGTCGACCGCTGAGGACGAGATGAGTACGATGATCTTCTCTG 1647
 Db 1909 AGCAACTTCCGCTGATGAGTGGTGGACGAGATGCGCATGATGACGTCTTATTTATT 1968
 QY 1648 GGGACTGCTAGGCTGTGCTCAAAATGCTATGCTCTCCAGCGAGGGGCTCAGCTGAA 1707
 Db 1969 GGGACGACACAGGAATTTGCTGAAGTATACAAATTTACAAACAAAGAGTGG 2028
 QY 1708 CCTGAGGAAGTGTTCGAGGAGCTCCAGGTGTTAAGTGGCAACACCTATACCGAA 1767
 Db 2029 ATGAGAGATGATCTTCTAGAGGAACTTCAATATTCAGATTCAGCCCTATCATTTCT 2088
 QY 1768 ATGAGATGCTCTGCAAAAGGCAATGCTATGCTGAGGCTGGCTGGGTGGCCAG 1827
 Db 2089 ATGGAATTTCTTCAAAAGACAAAGCTTTTACATTTGATACAGCTCTCTGCTGGCAAA 2448
 QY 1828 CTGCGGCTGACCAATGTAGATTCAGGCTGCTGTCAGAGTGTGCTGGCGCGG 1887
 Db 2149 GTGAGATTCATACATGCGACATGTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGA 2208
 QY 1888 GACCATATGCTGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
 Db 2209 GACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1941
 QY 1942 GCGAAGCGCGTTCGCGGCGAGACATGCGGAGCAACCTGCTGCTGCTGCTGCTGCTG 2268
 Db 2269 GCGAAGGAGGAGTTCGCGGAGCAAGCTTGGCATGCGCAAGCGCGCAAGAGTGTGTT 2338
 QY 2002 GCGCAGAGCGCGAAGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338
 Db 2339 GCGCAGAGCTTGTGAGAGCGCTTGGACAGGACGATGAGGCTGCTGCTGCTGCTGCTG 2061
 QY 2062 GAGCACAATGACACTTCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
 Db 2389 GAGGCAACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2121
 QY 2122 CTCTTCGAGAGCGCGAGGAGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
 Db 2449 TTTGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2181
 QY 2182 CACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2508
 Db 2509 AAGATGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2241
 QY 2242 TGCACCACTTGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
 Db 2569 TGCACAGACAGTAGAACAATTTTGTCTATCTGTGCTTAAATCACCCTTGGAGCTGCTG 2628
 QY 2302 GTGCGCTCAGCGTGGAGCACTTTCCTTCCTGCGAGC-----CAAGCCAGAG 2349
 Db 2629 GAGAGAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2688
 QY 2350 GAGCGCCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2409
 Db 2689 ATGCGCTGCGCTCCTTAAGCGGTATGTCTCAGGAGGAGCAAAACCGGTATCAAGGATTC 2748
 QY 2410 CTGACACTATTTGGCTTGGCAACTGCGCGCGGTGATGAGTACTGATGAGCGCTGCTG 2469
 Db 2749 TTGCACTACTATTTGGCTTGGCAACTGCGCGCGGTGATGAGTACTGATGAGCGCTGCTG 2808

QY 2470 TCGA 2473
 Db 2809 TGTGA 2812

RESULT 9

AA228470
ID AA228470 standard; DNA: 3988 BP.

AA228470;

05-JAN-2000 (first entry)

Mouse semaphorin H variant (Sema Hv) polynucleotide sequence.

Semaphorin H; Sema H; Sema Hv; collapsin; bone structure formation; metastasis; cancer; antibody; drug screen; ss.

Mus sp.

Key Location/Qualifiers
CDS 609..2943
FT /tag= a
FT /product= Sema Hv
FT /note= "Mouse semaphorin H variant"

W09947671-A2.
23-SEP-1999.

12-MAR-1999; 99WO-1B00495.

13-MAR-1998; 98US-0077997.

(LURA) LUKANIDIN E M.
(CHRL) CHRISTENSEN C R L.

Lukanidin EM, Christensen CRL;
WPI: 1999-590975/50.
P-PSDB: AA43091.

New polypeptides and polynucleotides, useful in diagnosis and treatment of metastatic cancer
Claim 1: Page 85-91; 95pp; English.

This is the mouse Semaphorin-H variant (Sema Hv) polynucleotide sequence (Sema Hv). The Semaphorin/Collapsin family of molecules are characterised by a unique and highly conserved motif, within a 500 amino acid semaphorin domain. Some semaphorins exhibit inhibitory or repulsive functions in a neuronal context, and functions in bone structure formation are also implicated. It is thought that Sema H plays a role in metastatic cancer. The invention uses the mouse Sema H gene in the diagnosis of determining the metastatic potential of cells, by detecting their expression in biological samples. Antibodies specific for Sema H, are also useful therapeutically in inhibiting Sema-H polypeptide activity and also be inhibited by inhibiting the polypeptides. Metastasis may polypeptide using e.g. a small molecule inhibitor or a Semaphorin-H ligand (or fragment). The polynucleotides can also be used to inhibit prevent metastasis of cancer cells. They can be used to detect and quantify Sema-H mRNA levels in cells. The polypeptides, fusion proteins, included in pharmaceutical compositions. The polynucleotides can be used to isolate similar sequences from other species and to produce mammalian cell lines and tumours with known metastatic potential, useful in anti-metastatic drug screening.

1288 CAGGCGCTTCGCCGTCTGTGTGTACCAATGGCAGACATCTGGAGGTTTTCACACGGGCC 1347
1618 AGAGGGCCATGCTGTATGTGTATTCACATGTCACAGTATCCGGAGAACCCCTTATATGGCCA 1677
1348 TTGGCCACAGGATGTGGGCTTCAGCACCACTGGGGGGCCATATGGGGGCAAGGTGCCCTTC 1407
1678 TATGCTCATAAAGAGGCCCTGAAATACACTGTGTACTATATGAAGAAAGATGCCCTTCAC 1737
1408 CCTGCCCTTGGCGTGGCCCCCAGCAAGATGACCGCACAGCCAGAGCGCCTTTGGCAGC 1467
1738 CCAAGGCTGCTTCTCTGTGTGCACAGCAAAAGTAAAC-----GGAGCCAGTATGGAAC 1788
1468 ACCAAGSACTACCCAGATGAGGTGTGCAGTTTGGCCAGGCCACCCCTCATGTTCG 1527
1789 AACCAAGATTATCCCGATGACGCCATCCGTTTCGCAAGATGCAATCCCTCATGTATCAG 1848
1528 CCTGTGCGGCTTCGCATGTGGCGCCCTGTCTTGTACAGACCCACCTTGGCCACAGCTTA 1587
1849 CCGATTAACCTGTTCAATAAAAAACCAATCTGTGTAAAAACAGATGAATATACACCTG 1908
1588 CACCAATCTGTGTGTGGACCGGTGTGAGSAGAGAGATGTGGACCTACGATGTCATTTCTG 1647
1909 AGCACACTTGCCGTGTATGTGGGTGAABCGAGSAGATGGCCAGTATACCTCTTATTTAT 1968
1648 GGGACATCATCAGAGGCTG---TGTCAAETCATGCTCTCCAGCCAGGGGCTCACT 1704
1969 GGCACACACACAGAGTATGTGCTGTGAATATTCACAAATTTACACCAAGAACAGAG 2028
1705 GAACCTGAGSAAATGCTGTGAGGAGCTCCAGGTGTTTAAGTGTCCACACTATACAC 1764
2029 TGGATGAGAGAGTCAATCTTAGAGAACTTCAATATTTCAAGGATCCAGCCCTATCAT 2088
1765 GAAATGAGATCTCTGTCAAAGAGCAATATGATATACGTGGGCTTCGCTGTGGGTGGCC 1824
2089 TCTATGAAATTTCTTCAAGAGACAAACACTTTACATTTGATCAGCCTGTGCTGTGGCA 2148
1825 CAGCTGGCGGTGCACCAATGTGAGCTTACGGCACTGCTGTGCGAGATGTGCTGCTGGCC 1884
2149 CAACTCAGATTCATCACTGTGAGATATTAAGGACAGTGTGTGCTGTGCTGTGCTGCT 2208
1885 CGGAGCCATCTGTGCTGTGGGATGTGGCTCTGTACCCACTACCGCCCA-----GC 1938
2209 CGAGACCCCTACTGTGGCTGTGGATGGCATATCTGCTCCAGGTATACCAACAGGTGCA 2268
1939 CTTGGCAGACGCGGTGTCCGCGGAGAGACATCCGGAGCGCAACCTGCTTCAGTGC 1998
2269 CACGAAAAGGAGGTTCCGCAAGCAGACCTTGGCATGGCAACGCCGCCCAACAGTGC 2338
1999 CTGGGCCAGGCGAGGAAAGAGCAGTGTGGACTTGTGGCAACCACTATGTCTACAGGC 2058
2339 TTTGACAGCAATTTGTTGGAGACGCGTTGGACAGGACTGAAGAGAGCTGTGATGGC 2388
2059 ACGGAGCAATATGACACTCTCTCGAGTGCCCGCCCAAGTCTCCCAAGCTGTGTCGCGC 2118
2389 ATAGAGACCAACATATCTTGTGGATACACCCCGCATATCATTAAGAGAAAGTATC 2448
2119 TGGCTCTTGAGGAGCGAGGAGTGTGGGGCTGTACACAGGTGAAGAGGAGAGAGAGTGC 2178
2449 TGGTTTCTACGAAGGAGACCGGACGTAAAGAAAGAGAGTGAACGATGACAGAGTT 2508
2179 TTGCAACAGGAGCGGGGCTCTGTCCGAGGCTTAGCCGTTTGTGATCGGCACTAC 2238
2509 GTCAAGATGTGATGTGGCTTCTCTTCCATGAGTACGCAATGTAGATGACGGAGCTAT 2568
2239 ACTGTACCACTGTGAGCATGTGCTTCTCCACATGTGTGTCGCTGTGCTGTGTTG 2298
2569 TTTTGGCAGCAGTAGAACAACAAATTTTGTCCATATCTGTGCGTAAATCATCTTGGAG 2628
2299 ATTGTGCTCACAGCTGTGCAACCTGTTCCTCCGAGC-----CAAGGCA 2346
2629 GTCGAAGAGCATTAAGTGTGAGGCGATGTTTATTAAGAGACATGAAGAGAAAGACATCAC 2688

1257 CGCGCTTACAGCAGCCGTCAGTGGCGTGTTCAGAGGCTTCCGCTCTGTGTACACAT 1316
 1258 TGGAGTGTACAGTTCAGTACATTTTCAGAGGATCCGCTGTGTATGTATGATCAT 1049
 1259 TGGAGTGTACAGTTCAGTACATTTTCAGAGGATCCGCTGTGTATGTATGATCAT 1376
 1317 GGCAGATCTGGAGAGTGTTCACGGGCTTGGCCACGAGATGGCCCTTACAGCA 1109
 1318 GAGTGTGTAGAGAGGATGTCTCTGTGTATGTATGTATGTATGTATGTATGTAT 1436
 1050 GAGTGTGTAGAGAGGATGTCTCTGTGTATGTATGTATGTATGTATGTATGTAT 1168
 1377 GTGGGGGCCCCATTTGGGGGAGGTGCTTCCCTGGCTGGCTGGCTGGCTGGCTGG 1168
 1378 GTGGGGGCCCCATTTGGGGGAGGTGCTTCCCTGGCTGGCTGGCTGGCTGGCTGG 1496
 1110 ATGGGTGCTTATCAGAGAGAGTCCCTATTCAGGCGGAGAGACTGTGCCAGCAAA 1217
 1437 GACCCGACAGCAGGAGGCTTTTGGCAGCAGCAGAGACTTACCAAGATGAGTGTCTCA 1217
 1438 GACCCGACAGCAGGAGGCTTTTGGCAGCAGCAGAGACTTACCAAGATGAGTGTCTCA 1556
 1169 -----CATTTGTGTTTGAATCTACAAAGAGCTTGTGTATGTATGTATGTAT 1277
 1497 GTTGGCCGAGCCAGCCCTTCATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1277
 1218 CTTTGCAGAGAGTATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1616
 1557 CTTTGCAGAGAGTATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1337
 1278 ACTGATCAAAAGAGTATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1676
 1617 AGAGATGGGAGCTATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1397
 1338 AGAGATGGGAGCTATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1736
 1677 CATGCTCTCAGAGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1457
 1398 AGTTCATATTCAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1796
 1737 GGTGTAAAGT 1517
 1458 AGT 1856
 1797 ATACGTGGGCTGT 1577
 1518 ATATATGT 1916
 1857 CATGCTCTCAGAGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1637
 1578 GAAAGGCTGT 1976
 1917 CTGTACCACTACCGCCAGCCTTGGAGAGCGCGGTCCGCGGAGAGATTCAGTGTGT 1694
 1638 ATGTCTCGCTATTTTCCCA---CTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2036
 1977 CGGCAACCTTCCCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1754
 1695 TGGAGATCCCACTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2096
 2037 GGCAGACCATGT 1814
 1755 TGAAGAGAAATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2156
 2097 GTTCCCAACCTGT 1874
 1815 GTCCAGAGAGAGCTGT 2216
 2157 GGTGAACAG 1934
 1875 GATCAGAGT 2276
 2217 CGGT 1994
 1935 ACAGAGAGATTCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2336
 2277 GGTCCGCTGT 2054
 1995 TCTTAAGTAACTGT

2337 GCCAAGCCAG 2390
 2055 TGTATGT 2450
 2391 GGCCTGT 2174
 2115 GGT 2478
 2451 GTACTGT 2202
 2175 GTTCTGT

RESULT 11
 AA087442 standard; cdna: 2601 BP.
 ID AA087442
 AC AA087442
 DT 21-NOV-1995 (first entry)
 XX
 DE Human semaphorin III cdna.
 XX
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
 KW varicella major virus; smallpox; semaphorin receptor binding activity;
 KW modulation; nerve cell growth; immune response; viral pathogenesis; ds.
 KW neurological disease; neuro-regeneration; oncological infection; ds.
 KW Homo sapiens.
 OS
 XX
 FH Key
 FT 16..2331
 FT CDS /tag= a
 FT /product= human semaphorin III
 XX
 PN W09507706-A.
 XX
 PD 23-MAR-1995.
 XX
 PF 13-SEP-1994: 94MO-US10151.
 XX
 PR 13-SEP-1993: 93US-0121713.
 XX
 PA (RESC) UNIV CALIFORNIA.
 XX
 PI Bentley DR, Goodman CS, Kolodkin AL, Mattes D;
 PI O'Connor T;
 XX
 DR WPI: 1995-131177/17.
 DR P-SDB: AAR1380.
 XX
 FT New class of semaphorin peptide(s) and polypeptide(s) - are
 FT potent modulators of nerve cell growth and regeneration
 XX
 PS Example 2: Page 60-63; 101pp; English.
 XX
 CC The sequence of the cdna encoding the human semaphorin III protein.
 CC The proteins encoded by the grasshopper semaphorin I (AA087443), Drosophila
 CC semaphorin III, vaccinia virus semaphorin IV (AA087445) or
 CC semaphorin I and II (AA087444-5), Tribolium semaphorin I (AA087446)
 CC varicella major (smallpox) virus semaphorin IV (AA087447), which retain
 CC to generate a series of binding activity. The semaphorin derived or
 CC semaphorin receptor binding peptides are potent modulators of nerve cell
 CC semaphorin receptor binding activity and viral pathogenesis. They can be used
 CC in diagnosis and treatment of neurological disease and
 CC neuro-regeneration, immune modulation and diagnosis and treatment of
 CC viral and oncological infection and diseases.
 CC
 CC Sequence 2601 BP; 809 A; 533 C; 593 G; 666 T; 0 other;
 SQ
 Query Match 18.1%; Score 474.8; DB 16; Length 2601;
 Best local similarity 53.9%; Pired. No. 8.4e-95;

7:

7:

7:

Page 20

KW	Human; immune; hematopoietic; immune/haematopoietic antigen; cancer
KW	cytostatic; gene therapy; vaccine; metastasis; ds.
XX	
XX	Homo sapiens.
OS	
PM	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01354.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0205467.
PR	28-JUN-2000; 2000US-0209486.
PR	30-JUN-2000; 2000US-0214886.
PR	07-JUL-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	11-JUL-2000; 2000US-0216880.
PR	14-JUL-2000; 2000US-0217487.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	14-AUG-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	18-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226686.
PR	22-AUG-2000; 2000US-0227182.
PR	23-AUG-2000; 2000US-0227009.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.
PR	08-SEP-2000; 2000US-0231242.
PR	08-SEP-2000; 2000US-0231243.
PR	08-SEP-2000; 2000US-0231244.
PR	08-SEP-2000; 2000US-0231413.
PR	08-SEP-2000; 2000US-0231414.
PR	08-SEP-2000; 2000US-0232080.
PR	12-SEP-2000; 2000US-0232081.
PR	14-SEP-2000; 2000US-0232081.
PR	14-SEP-2000; 2000US-0232397.
PR	14-SEP-2000; 2000US-0232398.
PR	14-SEP-2000; 2000US-0232399.
PR	14-SEP-2000; 2000US-0232400.
PR	14-SEP-2000; 2000US-0232401.
PR	14-SEP-2000; 2000US-0233063.
PR	14-SEP-2000; 2000US-0233064.
PR	14-SEP-2000; 2000US-0233065.
PR	21-SEP-2000; 2000US-0234223.
PR	21-SEP-2000; 2000US-0234224.

[illegible]

ID AAX03792 standard; cDNA: 2975 BP.
 AC AAX03792;
 XX
 XX
 XX 31-MAR-1999 (first entry)
 DE Human semaphorin E encoding cDNA clone BR5334.
 XX
 XX Human; semaphorin E; clone BR5334; nutritional; immune stimulating;
 KW vaccine; haematopoiesis regulating activity; tissue growth;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 XX gene therapy; ds.
 OS Homo sapiens.
 XX
 XX
 XX Key Location/Qualifiers
 FT CDS 245..2500
 FT /*tag= a
 XX
 XX W09853065-A1.
 XX
 XX 26-NOV-1998.
 XX
 XX 19-MAY-1998; 98MO-US10188.
 XX
 XX 18-MAY-1998; 980S-0080695.
 XX 19-MAY-1997; 970S-0858834.
 XX
 XX (GENM) GENETICS INST INC.
 XX
 XX Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER;
 PI McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
 DR WPI, 1999-059742/05.
 XX P-PSDB; AAM30617.
 XX
 XX New polypeptides encoding secreted human proteins - derived from a
 PT human foetal kidney cDNA library
 XX
 XX Claim 1; Page 43-44; 58pp; English.
 XX
 CC The present sequence encodes human semaphorin E from cDNA clone BR5334.
 CC Human semaphorin E polynucleotide sequences and protein sequences from
 CC the present invention, are predicted to have biological activities which
 CC would make them suitable for treating, preventing or ameliorating
 CC medical conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC adherin/tumour invasion suppressor activity, and tumour inhibition
 CC activity. The polynucleotide sequences are also stated to be useful for
 CC gene therapy. A host cell transfected with the polynucleotide sequence
 CC encoding human semaphorin E or its subfragments and variants is useful
 CC for recombinant production of the clone BR5334 related protein.
 XX
 XX Sequence 2975 BP; 887 A; 620 C; 665 G; 802 T; 1 other;
 QY
 Query Match 14.98; Score 391.6; DB 20; Length 2975;
 Best Local Similarity 53.18; Pred. No. 1.6e-76;
 Matches 1004; Conservative 1; Mismatches 855; Indels 30; Gaps 7;
 QY 394 GACTCTCTGTGCGCAACCGCTGTGACATCTTGTGGCCCCCGAGGCTCCCTGAACTTC 453
 DB 347 GAACTTGTGAGAAACCAAGACCTGTGATTAATTCAGCTTTTCCACATCTTTAGACTAC 406
 QY 454 CAGGCACTGTACCTAGTAGTACCGAGACGCGCTTTCTGGGTGGCGGAGCGGCTC 513
 DB 407 AGGATTTTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 466
 QY 514 TACTCTCTGCGCTGTGACCAAGCGATGCGGAGATGCTGTGGCGACCGGAG 573

DB 467 CTTTCCCTGATATTTAACAATATAAGTCAAGAGCTTTGAGTGTTCCTGGCCAGCATCT 526
 QY 574 CCAGACAGAGGAGGAGAGTGTGTCGAAAGAGAGATCTCTTTGACAGAGTGGCCGAC 633
 DB 527 ACAATCAAAAGTTGAAGATGCAAAATGCTGGCAAAATGCTGGCAACACAGCTGGGAG 586
 QY 634 TTGCTGGCGGTCACAGAGCTTCACACACCGGACCCAGCTGCTAGCTGTGGGAGGCGC 693
 DB 587 TTGCTGGCGGTCACAGAGCTTCACACACCGGACCCAGCTGCTAGCTGTGGGAGGCGC 693
 QY 694 TTGCGAGCCACCTGTGCGCTCATCACAGTTGGCCACCGTGGGAGCATGTCTCCACTG 753
 DB 647 TTGAGTCTGCTGTGCTACTTACTTGAACAGAGGAGAGATCAGAGGACCAAGTTTTCATG 706
 QY 754 GAGCTTGGCAGTGTGAAAGTGGCGGCGGCGGCTGCTCAGAGCGCGGCTGCTT 813
 DB 707 ATTGACTCCAAAGTGTGAATCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
 QY 814 GCGAGCCTTCAATAGACGGGAGGAGCTGTACAGGGGTCACAGCTGACTTCTGGGCGGA 873
 DB 767 GTGTCTGTATATGATCAATGAGGAGGAGCTTTCTCTGSAATGTATATAGATTTTATGGG 826
 QY 874 GAGGCCATGATCTTCCGAAGTGGAGGAGTCTGCGCAGCTGCTGCTGCTGCTGCTGCTG 933
 DB 827 GATGCTGCTATTTTTCGAAGTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 886
 QY 934 AGTCTCTGCG--ACGACCCCGGTTTGTGATGAGCCCGGATCCCTGAGAACTGTAC 990
 DB 887 TCCAAATGGCTAAGTGAACCTATGTGTGTATGATGATGATGATGATGATGATGATGAT 946
 QY 991 CAGACATGACAGAGTGTACTTCTTCTGAGAGAGGTCCTCCCTGCGGAGTGTGTC 1050
 DB 947 CCAATATGCTAAGTGTACTTCTTCTGAGAGAGGTCCTCCCTGCGGAGTGTGTC 1006
 QY 1051 TCGAACCATGTCACTGTGACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1110
 DB 1007 ACGAACAGATTCATTTCA--TGATTGCTCGAATATGTCTAATGACACTGTGTGACTG 1063
 QY 1111 CGGCTGTGTGTGAACAATAGAGAGACCTTCTCAAGGCGGCTGTGCTGCTGCTGCTG 1170
 DB 1064 CGTACCTGTGTGAACAATAGAGAGACCTTCTCAAGGCGGCTGTGCTGCTGCTGCTG 1123
 QY 1171 GCGCTGTGTGTGAACAATAGAGAGACCTTCTCAAGGCGGCTGTGCTGCTGCTGCTG 1230
 DB 1124 GATGAAGAGCGGCGGCAAAACACTTGTATGATGATGATGATGATGATGATGATGAT 1183
 QY 1231 AAGGCGGGAAGAGCTGTGAGTGTGACGCGCTGTGACAGACCGCTGAGGCGGCTGTG 1290
 DB 1184 GATTAACCGGAGAGACACTGATGTATGTGCTATTTTTCACATCACTGATGATGATGAT 1243
 QY 1291 GCGTGTGCGCTGTGTGTGTACCATGTGACAGATCTGAGAGTGTTCAGAGGCGGCTT 1350
 DB 1244 GGATTCACCGCTGTGTGTGTGTATCATTTATCTGATATGAGCTGTGTTATGAGGCTTT 1303
 QY 1351 GCCCAGCGAGATGGGCTGTGACGACCACTGTGAGGCGGCGGCGGCGGCGGCGGCGG 1410
 DB 1304 GCCCAAGAAAGAGGCGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1363
 QY 1411 CGCCCTGCGCTGTGCGGCGGCAAGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1470
 DB 1364 GCGCCTGTGAATCTGTCAAGA-----GAGCATTTTACACCCCAATATGGAACAC 1414
 QY 1471 AAGGACTTACCAAGATAGTGTGTCGAGTTTGGCGAGCGGCGGCGGCGGCGGCGGCGG 1530
 DB 1415 AAGGAGTTTCCAGATGATGTGTCTCACTTTTATTCGGAACACCTCTCATATGATACAT 1474
 QY 1531 GTGGCGCTTCGACATGCGCGCGGCTGTGCTGTGCAAGACCGACCTGTGGCGGCGGCGG 1590
 DB 1475 ATCTTCCCAATTCACAAAGGCTTTGATTTGTGTGATGTGATGATGATGATGATGAT 1534
 QY 1591 CAGATGCTGTGTGAGCGGCGGCGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1650
 DB 1535 AAGATAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594

XX	Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;	
PI	Merberg D, Treacy M, Spaulding V;	
XX	WPI: 2000-679620/66.	
DR	P-PSDB: AAB28379.	
XX	New monoclonal antibodies, useful for treating cancer and	
PI	immunodetection of secreted proteins which are in turn useful for	
PI	treating neurological, inflammatory, immune diseases and microbial	
PI	infections	
XX	Disclosure: pages 66-67; 75pp; English.	
PS		
XX	The present sequence is the coding sequence for human clone BR533.4. The	
CC	present sequence was isolated from a human foetal kidney cDNA library.	
CC	The BR533.4 protein is a secreted protein and can be used to treat a	
CC	number of conditions including various immune deficiencies and disorders,	
CC	microbial infections, autoimmune disorders, allergic reactions such as	
CC	asthma, respiratory problems, inflammation, myeloid or lymphoid cell	
CC	deficiencies, anaemias, burns, wound healing, ulcers, periodontal	
CC	disease, central and peripheral nervous system diseases and neuropathies,	
CC	lung or liver fibrosis and cancer.	
XX		
SQ	Sequence 2975 BP; 887 A; 620 C; 665 G; 802 T; 1 other;	
XX		
Query Match	14.98; Score 391.6; DB 21; Length 2975;	
	.Best Local Similarity 53.18; Pred. No.1.6e-76;	
	Matches 1004; Conservative 1; Mismatches 855; Indels 30; Gaps	
OY	394 GACCTCCCTGCTGGCAACCGCTCGCCATCTTCTGCGGCCCCAGGGCTCCCTGAAACCTC 453	
DB	347 GACCTTCGAGAAACCCAGACCTCTGGAATCTTACGCTTTCGCCACATCTTGAAGACTAC 406	
OY	454 CAGGCCATGTACCTAGATGAGTACCGAGACGCCCTCTTGGGTGGCTGGAGCCCTC 513	
DB	407 AAGATTTTATATGATGAGATGAGATGAGACCGGATATATGTGGAGACCAAGATCATTT 466	
OY	514 TACTCTTCGGCTGGACGAGCATGGCCAGATCCCGGAGGTCCTGGCCACCGGAG 573	
DB	467 CTTTCCCTGAATTTACAAATATAGTCAAGAAAGCTTTGAGTGTTCGGCCAGCATCT 526	
OY	574 CCAGGACAGAGGAGGAGTGTGTTCGAAAGGAGAGACTTTTGACAGAGTGGCCAC 633	
DB	527 ACATCAAAATGTGAAATTCACAAATGCTGGCAAGATCCACACAGCGCTGTGGAGAC 586	
OY	634 TTGCTGGGGGTGCTACAGCTTCACAAACCGGACCCACTGTACCTGTGGGACTGGGGCC 693	
DB	587 TTTTCCGTTATTCGACTTTCAATCGCACATTTGATGTGTGGAGTGGCGCT 646	
OY	694 TTCCAGGCCACTGTGCTTCATCCACAGTGGGCACGTTGGGAGCATGTGCTCACTTG 753	
DB	647 TTCACTCTGTCTACTTACTTACTTGAACAGAGGAGAGATCAAGAGACCAATTTTCATG 706	
OY	754 GAGCGTGGCACTGTGGAAGTGGCGGGGGGGTGGCCCTACAGAGCCGACCGCTCCCTT 813	
DB	707 ATTGACTCCAAGTGTGATCTGTGAAGAAAGAGCGCTCTTTCACCCCAACGTGACACG 766	
OY	814 GCCAGCACCTTCATAGACGGGAGACTGTGACAGGGGTCTCATCTGTGACTCTCTGGGCCGA 873	
DB	767 GGTGTGTGATGATCAATGAGAGGCTTTTCTCTGTGAATGTATATAGATTTCATGCGGACA 826	
OY	874 GAGGCCATGATCTTCCGAAGTGGAGTCTCTGGCCAGCTCTGCGTCTCCGACTGTGACAG 933	
DB	827 GATGTGCTATTTTTCGAAGTTTAAACCAAGAGATGGGTGAGACTGATCAACATATAT 886	
OY	934 AGTCTCTTG--ACGACCCCGGTTTGTGATGGCGCGCCGGATCCCTGAGACTGTGAC 990	
DB	887 TCCAAATGGCTAAGTAAACCTATGTTGTGATGATGACATGATCATCCAGATGGTACTAT 946	
OY	991 CAGGACAATACAGAGTGTACTTCTCTCTCTGGAGAGGTTCCCTCCGCGCCATGTTGGC 1050	
DB	947 CCAATATGCTAAGGTGCTACTTCTTCTCAAGAAACCACTGACTGACATATACGAGAGC 1006	

Accession	Sequence	Position
Qy 2131	AGCGCAGGGGATGAGGGGCCGTGACCAAGTGAACGCGAGCGAGCTTTGCACAGCGAG	2190
Db 2063	A---AAGCAAAAGACAGAGAGAAAGAGTTAACTTAATGAACGATTAATGCCACTTCA	2119
Qy 2191	CGGGGGCGCTGTCCGCGAGGCTTAAGCCGTTTGATGCGGGACCTAGACCTGCGACACT	2250
Db 2120	CAGGACTCTCTGATCGGCTCTGTTCAGGGTTCTGACCAAGACTTATACATCGCATTTGCT	2179
Qy 2251	CTGAGCATGAGCTTCCACAGACTGTGGTGC	2280
Db 2180	ACAGAAATAGTTCAAGCGAGACCAAGCC	2209

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 10:48:08 ; Search time 2035 Seconds
(without alignments)
17429.974 Million cell updates/sec

Title: US-09-813-290-1
Perfect score: 2628
Sequence: 1 atggcgtgtgccttagctg.....ggagagtgagagccagtag 2628

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estcl:*
9: gb_estcl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466.8	17.8	3436	11 AK014333	AK014333 Mus muscu
2	418.4	15.9	420	9 AW803853	AW803853 IL2-DM008
3	369.4	14.1	647	10 BI067436	BI067436 Pgf1n.PK0
4	357.6	13.6	364	9 A1750874	A1750874 cm06b10.x
5	320.2	12.2	576	10 BI067451	BI067451 Pgf1n.PK0
6	308.2	11.7	436	10 BM287610	BM287610 Pgf1n.PK0
7	280.4	10.7	645	10 BI066459	BI066459 Pgf1n.PK0
8	264.2	10.1	785	10 BI069389	BI069389 Pgf1n.PK0
9	255.2	9.7	609	10 BI066327	BI066327 Pgf1n.PK0
10	250.2	9.5	990	10 BI552258	BI552258 603299702
11	249.6	9.3	662	10 BI557438	BI557438 Pgf1n.PK0
12	243.8	9.0	920	9 A1541401	A1541401 A1541401
13	237.4	8.9	786	9 A1584324	A1584324 A1584324
14	235.2	8.9	754	9 A1584625	A1584625 A1584625
15	231.2	8.8	823	10 BI645606	BI645606 603275104
16	219	8.3	624	9 AW328571	AW328571 ds03d12.x
17	217	8.3	735	9 AU140116	AU140116 AU140116

18	215	8.2	737	10 BI646227	BI646227 603276957
19	212.6	8.1	343	10 H26438	H26438 y151b03.t1
20	210.2	8.0	626	9 AU137591	AU137591 AU137591
21	207.2	7.9	797	10 BI100405	BI100405 602885885
22	205.6	7.8	748	9 AU136052	AU136052 AU136052
23	203.8	7.8	1121	10 BM477901	BM477901 AGENCOURT
24	202.6	7.7	564	9 AW659026	AW659026 95939 MAR
25	202.6	7.7	678	10 BI151058	BI151058 602916885
26	200.2	7.6	674	9 AU138132	AU138132 AU138132
27	197.8	7.5	690	9 AU137448	AU137448 AU137448
28	197.6	7.5	728	9 AU137554	AU137554 AU137554
29	193	7.3	629	9 BI16052	BI16052 BI16052
30	192.6	7.2	778	10 BG184444	BG184444 RST3370 A
31	189.4	7.2	721	9 AW742133	AW742133 up52f06.Y
32	188.2	7.1	594	10 BG009731	BG009731 OVI-CN031
33	186.8	7.1	594	10 BI065228	BI065228 Pgf1n.PK0
34	186.6	7.1	637	9 AU135310	AU135310 AU135310
35	186	7.1	891	9 AU151337	AU151337 AU151337
36	183.6	7.0	540	9 BE017963	BE017963 b73h10.Y
37	183.4	7.0	1095	10 BM473502	BM473502 AGENCOURT
38	180.2	6.9	557	9 AU137539	AU137539 AU137539
39	179	6.8	557	9 AW957332	AW957332 EST369522
40	178.2	6.8	494	9 AW957332	AW957332 RST34801
41	177.8	6.8	774	10 BG215143	BG215143 2M0275K21
42	174.8	6.7	434	12 AZ991314	AZ991314 IL3-HT061
43	169.4	6.4	468	10 BE710200	BE710200 IL3-HT061
44	169.2	6.4	568	10 BF93458	BF93458 OV0-CN021
45	169.2	6.4	602	10 BM425751	BM425751 Pgf1c.PK0

ALIGNMENTS

RESULT 1	AK014333	3436 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK014333				
DEFINITION	AK014333				
	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length				
	enriched library, clone:3230402E04;sem domain, immunoglobulin				
	domain (Ig), short basic domain, secreted, (semaphorin) 3B, full				
	insert sequence.				
ACCESSION	AK014333				
VERSION	AK014333.1	GI:12852112			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (Strain: C57BL/6J) 14, 17 days embryo head cDNA to				
	mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
	clone:3230402E04.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	Carninci, P., and Hayashizaki, Y.				
AUTHORS	1 (sites)				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,				
AUTHORS	2 (sites)				
	Carninci, P., Itoh, M., Konno, H., Okazaki, Y., Mutamatsu, M., and Hayashizaki, Y.				
	Normalization and subtraction of cap-trapper-selected cDNAs to				
	prepare full-length cDNA libraries for rapid discovery of new genes				
	Genome Res. 10 (10), 1617-1630 (2000)				
TITLE	3 (sites)				
JOURNAL	Shibata, K., Itoh, M., Nagao, S., Sasaki, N., Carninci, P.,				
MEDLINE	11042159				
PUBMED	20499374				
REFERENCE	Shibata, K., Itoh, M., Nagao, S., Sasaki, N., Carninci, P.,				
AUTHORS	3 (sites)				
	Kono, H., Akiyama, D., Nishitani, T., Tashiro, H., Itoh, M.,				
	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				
	Yonekura, S., Inoue, K., Togawa, Y., Iwata, M., Ohtsuka, E., Watabiki, M.,				
	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.				
	RIKEN integrated sequence analysis (RISA) system -384-format				
	sequencing pipeline with 384 multicapillary sequencer				

JOURNAL	GENOME RES.	10 (11), 1757-1771 (2000)
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4 (sites)	
AUTHORS	The RIKEN genome Exploration Research Group Phase II Team and the FANTOM Consortium	
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 665-690 (2001)	
REFERENCE	5 (bases 1 to 3436)	
AUTHORS	Adachi, J., Aizawa, K., Akabira, S., et al.	

TITLE	Direct Submission	Submitted (10-Jul-2000)	Webb's
JOURNAL			

COMMENT
Please visit our web site (http://www.)

FEATURES	source	end	start
Location/Qualifiers	1. .3436	37	3

gene	accession	embryo	cdna library
gene	1. .3436	14, 17 days	mouse cdna library
misc_feature	1. .3436		

BASE COUNT	SECRETED, (semaphorin) 3B ^a	SECRETED, (semaphorin) 3B ^a	SECRETED, (semaphorin) 3B ^a
754 a	942 c	1001 g	739 t

Query Match	17.8%	Score 466.8;	DB 11;	Length 3436.
Best Local Similarity	56.0%;	Pred. No. 2.1e-83;		
Matches 113;	Conservative 1;	Mismatches 768;	Indels 106;	Gaps 8.

[illegible]


```

Db 60 CACCCAGAGAGCGCTTTGGCAGACCAAGAGACTACCCAGATGAGTCTCAGTTGCC 1
RESULT 3
LOCUS B1067436 647 bp mRNA linear EST 15-JUN-2001
DEFINITION pfln.pk012.d10 normalized chicken fat cDNA library gallus
collapsin-5 [gallus gallus]g, mRNA sequence.
ACCESSION B1067436
VERSION B1067436.1 GI:14474958
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
REFERENCE 1 (bases 1 to 647)
AUTHORS Cogburn,L.A., Morgan,R.W. and Burnside,J.
JOURNAL Chicken ESTs from fat
COMMENT Unpublished (2001)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2832
Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES
source Location/Qualifiers
1..647
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pfln.pk012.d10"
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMDH10B"
/notes="vector: pSPOR1"
195 g 123 t 12 others
BASE COUNT 119 a 198 c 195 g 123 t
ORIGIN
Query Match 14.1%; Score 369.4; DB 10; Length 647;
Best Local Similarity 73.8%; Pred. No. 4.1e-64;
Matches 474; Conservative 0; Mismatches 167; Indels 1; Gaps 1;
QY 1049 GCTGACACCATGTCTACTGTGACCGCGGCGCGCTGTGCTGATGATGATGCGGCGC 1108
Db 3 CCAGAGAGAGCGCTTTGGCAGACCAAGAGACTACCCAGATGAGTCTCAGTTGCC 1
QY 1109 AGCGGAGTCTGTGTAACAATGAGACACTTCTCAAGCGCAGCGTGTCTGCTGCTGC 62
Db 63 NNAGGAGTCTGTGTAACAATGAGACACTTCTCAAGCGCAGCGTGTCTGCTGCTGC 1168
QY 1169 CCGGCGCTGTGTGCGAGAGACCACTTGTGACAGCTAGAGATGTCTGCTGCTGCTGC 122
Db 123 CCGGCTGTGTGCGAGAGACCACTTGTGACAGCTAGAGATGTCTGCTGCTGCTGC 1228
QY 1229 CCAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
Db 183 CCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1288
QY 1289 AGCGGAGTCTGTGTAACAATGAGACACTTCTCAAGCGCAGCGTGTCTGCTGCTGC 242
Db 243 AGCGGAGTCTGTGTAACAATGAGACACTTCTCAAGCGCAGCGTGTCTGCTGCTGC 1348
QY 1349 TTGCGCCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
Db 303 TTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
QY 1409 CTGCGCGTGTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
Db 363 CCGCGCGCGTGTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1468

```

```

QY 1469 CCAGAGATACCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527
Db 423 CCAGAGAGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 1528 CCTGTGCGCGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
Db 483 GCTGTGCGCGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1587
QY 1588 CACCAAGATCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
Db 543 CTCCAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1647
QY 1648 GGGACTGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1689
Db 603 NNAGCGATGACAGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 644
RESULT 4
LOCUS A1750874 364 bp mRNA linear EST 22-JUN-1999
DEFINITION clone.h10.x1 Normal Human Trabecular Bone Cells Homo sapiens
ACCESSION A1750874
VERSION A1750874
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 364)
AUTHORS Jia,L.B., Young,M.F., Touchman,J.W., Boufford,G.G.,
Beckstrom-Stenbery,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
P.G., Hotchkiss,R.N. and Francomano,C.A.
JOURNAL The Skeletal Genome Anatomy Project
COMMENT Unpublished (1997)
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 06 row: b column: 10
Seq primer: -21M13 forward primer (AB1).
FEATURES
source Location/Qualifiers
1..364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NHTBC.cn06b10"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/notes="Organ: Hip; Vector: pBluescript; Site:1; Ecort"
BASE COUNT 63 a 121 c 117 g 63 t
ORIGIN
Query Match 13.6%; Score 357.6; DB 9; Length 364;
Best Local Similarity 99.7%; Pred. No. 7.7e-62;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2040 AGCCACATGCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2099
Db 7 AGCCACATGCTTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 66
QY 2100 TCCCGAGCTGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGT 2159
Db 67 TCCCGAGCTGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGT 126
QY 2160 GAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2219

```

Accession	Sequence	Position
Db	127 GAAGACGAGACCGAGGTCTTGCACACGAGACGGGGGCTCTGTCCGACGGTTAGCCG	186
Db	1279	
Oy	2220 TTTCGATGCGGACCTACACCTGCACCACTGTGAGCATGGCTTCTCCAGACTGGT	2279
Oy	187 TTTCGATGCGGACCTACACCTGCACCACTGTGAGCATGGCTTCTCCAGATGGT	246
Db	2280 TTTCGATGCGGACCTACACCTGCACCACTGTGAGCATGGCTTCTCCAGACTGGT	2339
Oy	2280 CGGCTGGCTCTGTGGTGTGTTGTGGCTTCACAGCTGAGACCACTGTTCCTCCGAGCC	306
Db	247 CGGCTGGCTCTGTGGTGTGTTGTGGCTTCACAGCTGAGACCACTGTTCCTCCGAGCC	2397
Oy	2340 AAAGCCAGAGAGCCCCCAGCCCGGGGAGGCTGGCTTCACCCCAACCCAAAGGCTGG	364
Db	307 AAAGCCAGAGAGCCCCCAGCCCGGGGAGGCTGGCTTCACCCCAACCCAAAGGCTGG	

[illegible]

ACCESSION BI067431 GI:1447973
VERSION BI067431.1
KEYWORDS
SOURCE chicken.
ORGANISM *Gallus gallus*
*embryoto; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Enph; Galliformes; Phasianidae;*

REFERENCE AUTHORS TITLE JOURNAL COMMENT	1 (bases 1 to 27)
Cogburn, L. A., Morgan, R. W. and Burnside, Chicken ESTs from fat unpublished (2001)	
Contact: Larry A. Cogburn University of Delaware 101 N. 19717, USA	

Townsend Hall, Newark, DE 19713
Tel: 302-831-1335
Fax: 302-831-2822
Email: copburn@udel.edu www.chickest.udel.edu
Location/Qualifiers

```

FEATURES
source
1.576
/organism="Gallus gallus"
0.031"

```

```

/db_xref="taxon:9033"
/clone="pf1n.pK012.dg" chicken fat cDNA library"
/clone_lib="normalized
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMDH10B"
/molecule="vector: pSPORI1" 104 t
161 g 26 others

```

BASE COUNT	DB	Length
ORIGIN	10;	576;
Score	320.2;	

Query Match	12.2%	Score 220	Score 2.9e-54	Indels	1	Gaps	1
Best Local Similarity	72.6%	Pred. No.	0	Mismatches	154		
Matches	411	Conservative					

[illegible]

3 GCAGGAGCACACATCGTCAGCCCGGTCGCGTGC 1168

[illegible]

D6
63 AGAGGGGACCTGTTTCCGTCGTGTTGC 1220

1169 CCGGCGCTGCTGCCGAGACCCTTTCACACGTAGAGATGTTCTCCTGCTGAGA 182

123 CCGGTCCTGGCGGCATGATACCACTTTTATGAGCTGGAGCATGTCTTTTTCGACGAC: 1288

1229 CCAAGCGGGAGAGACCTCGAGGTGTACGGCGCTGTACGACACCGTACACCTGCTTTTC 242

D_b

183 CCAAGGATGGGAAGACCACGAGACTCTTAGCAGTCTTTAGGCGA...

OY	1289	AGGCTGTGGCCCTGTGTGTACCATGTGGAGACATCTGGAGGTTTCAACGGGCCCT	1348
Dd	243	AGGGTCGCCCGTCTGTGTCTACCGCATTGGCCGACATCCGTAAGTCTTAAAGGGGCCT	302
OY	1349	TTGCCACCGAGATGGGGCTCAGCACCAAGTGGGGCCCTAATGGGGACAAGTGCCCTTCC	1408
Dd	303	TTGGCCACCGGAGACGCCCTCTCACCACTGGGGAGACTACGAGGGCCGTGTGCCCTAAC	362
OY	1409	CTCGCCCTTG-CCGTGTGCCACGAGATGACCACGACCCAGACGGCGCTTTTGGCAGC	422
Dd	363	CGGCCCGGNNNGTGTGTCCAGACAAACCACCAACAGGCCCGGAGGGCAATACGACAGC	422
OY	1468	ACCAAGGACTACCCAGATGAGTGTCTGAGTTGGCCGAGCCCAACCCCCTCATGTTCTGG	1527
Dd	423	ACCAAGGACTNNNNNNGGAGGTGTGACTTTCGCCGCGCTCAACCCCCATGNTNNNGG	482
OY	1528	CCGTGTGGGCTTCGACATGAGCGCCCTGTCTCTTGACGAAGCCCATCGGCCAGACGTA	1587
Dd	483	GCTGTGGCCACGAGCATCACCGCCCGCTGCACNNNTGAAGACTGACCTTCACACAGCGCCTG	542
OY	1588	CACCAGATCGTGTGGAGACGGCGGTGA	1613
Dd	543	CTCAACTNNNNNTGACCGCTGTAGA	568

RESULT 6	436 bp	mRNA	linear	EST 28-DEC-2001
BM287610	5' /	CDNA	sequence.	
BM287610				

LOCUS	DEFINITION	ACCESSION
5283378	MARC	3BOV BOS LA
BM287610		
BM287610.1		GI:17996636

VERSION EST. taurus
KEYWORDS COW.
SOURCE taurus

ORGANISM	REFERENCE
Bos taurus	1 (bases 1 to 436)
Eukaryota	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Metazoa	1 (bases 1 to 436)
Choriata	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Cetartiodactyla	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Ruminantia	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Perocera	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Bovidae	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Bos	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett
Authors	Freking, B. A., Roberts, A. J., Stone, R. T., Smith, T. P. L., Grosse, W. M., Fairweather, S. C., Bennett

TITLE
Sequence evaluation of four pooled-tissue normalized bovine CDNA
libraries and construction of a gene index for cattle
Cattle, 11 (4), 626-630 (2001)

JOURNAL
MEDLINE
COMMENT
Genome Res. 11: 1077-1082 (2001)
21180013
Contact: Smith TPL
Animal Research Center
USDA, ARS, US Meat Animal Res. Station
NE 68993-30166, USA

PO Box 166, Clay Center, KS 68926
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Bases called and alt. trimmed with phred
Single pass sequencing. Bases called with the -mnscore 18
V0.980904.e. Vector identified by cross_match with the
-crossmatch 12 options.

and -minase-
PCR primers
FORWARD: AGGAACACCTCTGACCAT
REVERSE: GTTTCGCCAGTCACGAC
BACKWARD: GTTTCGCCAGTCACGAC
plate: 137 row: N column: 23
Seq primer: ATTGAGTGTCACTATAG
location/Qualifiers

```

FEATURES
source
1. .436
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1b="MARC 3B0V"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SpOx6; Site.1: XbaI; Site.2: XhoI;
library made from pooled tissue from marrow, alveolar
macrophage; ovary, fetal semitendinosus muscle, and fetal
longissimus muscle." 92 +

```

BASE COUNT	62 a	161 c	131 g	02


```

1. .785
/organism="Mus musculus"
"nvd/N"

```

FEATURES	SOURCE
Location/Qualifiers	
1. .785	
/organism="Mus musculus"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/db_xref="IMAGE:5355609"	
/clone_1fb="NCI CGAP Mam6"	
/clone_1fb="NCI CGAP Mam6"	
/sex="female, virgin"	
/tissue_type="infiltrating ductal carcinoma"	
/dev_stage="5 months"	
/lab_host="DH10b"	
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: oligo dr. Library constructed by life technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"	
161 a 239 c 200 g 185 t	
10.1% Score 264.2; DB:10; Length 785;	

[illegible]

QY	1842	ATG	1844
		11	
Db	774	CTG	776

RESULT 9	609 bp	mpna	linear	EST 15-JUN-2001
BI066327			cdna library	Gallus gallus
LOCUS				
BI066327			chicken fat	
			cdna	
			g119910362	
			ptgln.pk007.m5	
			normal	
			g1111449656	
			ptgln.pk007.m5.5	
			similar	
			g1111449656	
			cdna	
			clone	
			ptgln.pk007.m5.2	
			semaphorin	
			sem2	
			homo sapiens	
			dbj	
			U05498.1	
			g11114498132.1	
DEFINITION				

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
ref NP_064448.1	semaphorin sem2	Homo sapiens	mRNA sequence.	
ref XP_003233.1	semaphorin sem2	Homo sapiens		
(AB029496)	semaphorin sem2			
BI066327				
BI066327.1	GI:14473849			
EST.				
Chicken.				
Gallus gallus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Neornithes; Neognathae; Galliformes; Phasianidae;				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Archosauria: Aves, 1 (bases 1 to 609)	Phasianidae, Gallus.			
Cogburn L.A., Morgan, R.W.				
Unpublished (2001)				
University of Delaware				
Townsend Hall, Newark, DE 19717, USA				
Tel: 302-831-1335				
Fax: 302-831-2822				
Email: cogburn@udel.edu				
location/Qualifiers				
www.chickest.udel.edu				

FEATURES

Source	Accession	Score	DB	Length
/organism="taxon:9031"		255.2	10	609
/db_xref="pftn:ph007.m15"				
/clone="pftn:ph007.m15"				
/clone_1b="normalized chicken fat cDNA library"				
/sex="Male and female"				
/tissue-type="fat"				
/lab_host="E.coli EMDH109"				
/note="Vector: pSPOR1"				
	91 a	210 c	193 g	112 t
				3 others

Query Match	Similarity	97.9%	Pred. No. 3.4e-41	Indels	Gaps
Best Local	conservative	67.9%	0	Mismatches 176	
Matches	385				
OY	354	CCCCGCCCCAGAGGAGCCCGCTGCGGCTCTCCCTACGAGACCTCTCTCTGCGCCACCG	413		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
		CGCGGGAGAGACGCCCGCCGCGCTCCGCGCTGCTCTACCTCGGACCTCTGGGTATGACCG	102		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	414	CTCTGCCATCTTTCTGGGGCCCGCAGGGCTCTCTGATACCTCAGGCCATGTACTAGTATGA	473		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	103	CTCCCTCTCTTTCTTTGGCCACGCGGGCTTCCTGGGCTTCCTGGCTCTCCCTATCCGGATGA	162		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	474	GTACGAGACGCCCTCTTTCTGGGTGGCGCTGACGCCCTCTATCTCTCTGGCGGTGACCA	533		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	163	GTACCGCTGACCGCGCTTCTCATGTGGGGCAGAGATGTGCTCTACTCCCTCGCTGGACAG	222		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	534	GCGCATGGCCAGATCCCGCGGAGTCTGTGGACACCGCACCCAGACAGAGGAGAGAGTGTG	593		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	223	GGCACGCGCCGACACCCAGAGATCTACTGGCCGCCACTCCCTGGACAGAGAGAGATGTG	282		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	594	TGTTGGAAGGGAAGAGATCTTTGACAGAGTGGGCCCACTTCTGTGGGCTGTCTACGCC	653		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	283	TTTTTGGAAAGGGAAGAGACCCAGACCAACCGACTGTGTCCAACTACTCTCCGCTGGTGAAGCC	342		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	654	TTCACACCGGACCCACCGCTGAGCTGTGGACCTGGGCGGCTTTCACAGCCACCTGTGCGCT	713		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			
	343	CTACACAGACACACACTCTGCTGGCTGTGGGAGGAGGAGTGTCTCCACCCCGCTGTCCACTT	402		
		1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111			

Thu Oct 10 09:35:02 2002

us-09-813-290-1.rst

QY 714 CATCAGATTGGACCGTGGGACCATGTG-----CTCCACTGAGAGCTTGACAGTGTG 768
Db 403 CATCTACGTGGGGACACCGGCGTGAACCCAGCATGCTTCACTGAGCTTGATGACAG 462
QY 769 GAAATGTGGCGGGCGGCTGCTGCTGACAGAGCCGCGCTGCTGCTGCTGCTGCTGCTG 462
Db 463 GAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 462
QY 828 AGAGCGAGAGTGTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
Db 523 CGTGGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
QY 888 CCGAAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
Db 583 CCGAGCAGCAGCGGAGCGCGCGCGCGCT 914
583 CCGAGCAGCAGCGGAGCGCGCGCGCGCT 609

RESULT 10
LOCUS B1652258
DEFINITION 603299702F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5340217 5'
ACCESSION B1652258
VERSION B1652258.1 GI:15566494
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 990)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@email.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1864 Row: C Column: 02
High quality sequence stop: 892.
Location/Qualifiers
1. 990
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="5340217"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORE; Site: 1; NCI;
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigator's providing
samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999). Note: this is a NCI-CGAP Library."

FEATURES
Source
Location/Qualifiers
1. 990
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="5340217"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORE; Site: 1; NCI;
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigator's providing
samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999). Note: this is a NCI-CGAP Library."

Query Match 9.5%; Score 250.2; DB 10; Length 990;
Best Local Similarity 59.6%; Pred. No. 4; Le-40;
Matches 514; Conservative 0; Mismatches 338; Indels 11; Gaps 5;
QY 577 GGACAGAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
Db 71 GATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
QY 637 GTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
Db 131 GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 130
131 GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 130

BASE COUNT
ORIGIN 194 a 285 c 292 g 219 t
Query Match 9.5%; Score 250.2; DB 10; Length 990;
Best Local Similarity 59.6%; Pred. No. 4; Le-40;
Matches 514; Conservative 0; Mismatches 338; Indels 11; Gaps 5;
QY 577 GGACAGAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
Db 71 GATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
QY 637 GTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
Db 131 GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 130
131 GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 130

QY 697 CAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
Db 191 CAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
QY 754 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
Db 251 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
QY 814 GCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
Db 311 GCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 310
QY 874 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873
Db 371 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 370
QY 934 AGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 933
Db 431 AGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 430
QY 991 CAGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
Db 491 CAGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
QY 1051 TCGAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
Db 551 TCGAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 550
QY 1111 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1110
Db 611 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 610
QY 1171 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1170
Db 671 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670
QY 1231 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
Db 728 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 727
QY 1291 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
Db 788 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 787
QY 1350 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1349
Db 848 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
QY 1410 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1409
Db 908 TCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 907

RESULT 11
LOCUS B1557438
DEFINITION 603238962F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292039 5'
ACCESSION B1557438
VERSION B1557438.1 GI:15444752
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 662)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@email.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1M11738 row: k column: 16
 High quality sequence stop: 660.

FEATURES

source

1..662
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5292039"
 /clone_lib="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH10B"

Location/Qualifiers

Site_1: Noli;
 Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT.
 Average insert 2 kb. Library constructed by Life
 Technologies, catalog #12017-018. Investigators providing
 samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
 for transgenic model: Xu et al., Nature Genetics 22, 37-43
 (1999). Note: this is a NIH_CGAP Library."

BASE COUNT 134 a 209 c 168 g 151 t

ORIGIN

Query Match 9.5%; Score 249.6; DB 10; Length 662;
 Best Local Similarity 62.0%; Pred. No. 4.6e-40;
 Matches 418; Conservative 0; Mismatches 244; Indels 12; Gaps 1;

QY 1251 GGTGTAGCGCTTGTAGACACCGTCAGTCGCGTTCAGGGCTTCGCCGTCTGTGTGTA 1310
 Db 1 GCTGTATGCTGTCTCTCCACCTCCAGTGTGCTTCCAGGGCTTCGCTGTGTGTA 60
 QY 1311 CCACATGGCAGACATCGGAGGTTTTCACAGGGCTTTCGCCACCGAGATGGGCTTCA 1370
 Db 61 CAGATGAGACATGTGGCGCGAGGCTTCTTGAGACCTTTGCTCACAAAGAGGGCTTCA 120
 QY 1371 GCACCATGGGGGCGCTTATGGGGGCAAGGTCCTTCGCGCTGGCGCTGTGCCGCA 1430
 Db 121 ACACCATGGGCTGTCTTCAAGGGTCTGTGCTCCCTTACCAAGACCTGGCATGTGCCCA 180
 QY 1431 CAAGATGACCGCAGCAGCAGCGCTTTTGGCAGCAGCAAGCAGCAATCCAGATGAGT 1490
 Db 181 CAGA-----CCTTGGCAGCTTCAAGCTTCCACCAAGAGACTTCCAGATGAGT 228
 QY 1491 GCTGAGTTTGGCCGAGCCACCCCTCATGTCTTGGCTGTGGGCTTGCATGGCCG 1550
 Db 229 TATCCAGTTTGTGCGAACCACCTCTCATGTATCAACCAAGTCCGCTTGGGGGGCG 288
 QY 1551 CCGTGCCTTGTCAAGACCCACCTGGCCAGCAGCTACACCAAGATCGTGTGACCGGT 1610
 Db 289 CCGTCTTCTTCAAGTGGAGCTGTGATACACCTTCAACCAATGCGCGAGACGAGT 348
 QY 1611 GGAGCAGAGATGGACCTTATGATGATTTTCTGGGAGCTACTAGGGTCTGTGCT 1670
 Db 349 AGCAGCTGCGGATGACACTATGATTTCTTATGAGTACAGATGGGCGACAGT 408
 QY 1671 CAAGTCTATCCCTTCCAGGAGGGGCTCAGTGAACCTGAGAGATGTTCTGAGAGA 1730
 Db 409 GAATGATGTCAGTCCCAAGAGCAGCCGCAATATTCGTGAGAGATTTCTCTGAGAGA 468
 QY 1731 GCTCAGATGTTTAAGGCAACACCTATACCAAGATGAGATCTCTGTCAAAAGGA 1790
 Db 469 GCTGAGAGTTCGAGGACTCTGCGGTATATCAGCAGATGCAATCTCTTAAAGGA 528
 QY 1791 AATGCTATACGTGGGCTCTGGCTGGTGTGGCCAGCTGGCGCTGCAACATGTGAGAC 1850
 Db 529 ACAACTCTACGTAGATTCGCGAGCGCAGGAGTGGCCAGATGCTTGCATCTCTCAGTGC 588
 QY 1851 TTACGAGCTGCTGTGAGAGTGTGCTGGCGGAGCCATCTGCTGGCTGGAGATG 1910
 Db 589 CTTAGGCGCTGCTGGCAGAAATGCTGTGGCCGCTGATCTTATGCGCTGGGATGG 648

QY 1911 TGCCCTCCGTACCC 1924
 Db 649 ATCAGCTTGACAC 662

RESULT 12

AL541401

LOCUS

DEFINITION

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

AL541401

FEATURES

source

1..920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSDB005YL11"
 /clone_lib="LTI_FLO02_PL1"
 /lab_host="DH10B"

Location/Qualifiers

Site_1: placent; Vector: pCMVSPORT 6; 1st strand
 /note="Organ: placent; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies, Contact: Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax: (1) 301 610 8371 Email:
 fliang@lifetech.com URL:
<http://fulllength.invitrogen.com>"

BASE COUNT 170 a 276 c 306 g 167 t

ORIGIN

Query Match 9.3%; Score 243.8; DB 9; Length 920;
 Best Local Similarity 60.0%; Pred. No. 7.6e-39;
 Matches 495; Conservative 1; Mismatches 318; Indels 11; Gaps 5;

QY 352 AGCCCGGCCCCAGTGTGCGCGCTTCTTACCGAGACCTCTGTGCGCAC 411
 Db 100 AGTCCGCGCCCCAGTGTGCGCGCTTCTTACCGAGACCTCTGTGCGCAC 156
 QY 412 CGCTCTCCAGACTTCAAGCTGTGAGCGAATCTGCTTACGAGCCCTTGTGAGAT 216
 Db 157 CATGCTCCAGACTTCAAGCTGTGAGCGAATCTGCTTACGAGCCCTTGTGAGAT 216
 QY 472 GAGTACGAGACCGCTTCTTGGGGGCTTGGGAGCCCTTACTCTTGGGGGCTTAC 531
 Db 472 GAGTACGAGACCGCTTCTTGGGGGCTTGGGAGCCCTTACTCTTGGGGGCTTAC 276
 QY 217 GAGGAGCTGTGAGCGCTGTTGTGGGCGGAGAACATGTGCTTCCCACTGAGAC 590
 Db 532 CAGGATGCGCAG-ATCCCGGAGGCTGTGCGGACCGCAGCAGGAGAGAGAGGA 336
 QY 277 AACATCAGCTTACCGGCGGCGCAAGAGCTGGCTGCGCGCCCTGTGGAATGCGAGAGA 650
 Db 591 GTGTGTTGAAGGAGAGAGATCTTTGACAGATGCGCAGCTTCTTGGGGGCTTAC 396
 QY 337 GTGCAACTGCGGAGAGAGAGATGTGATGAGAGCTGTAAGTCTGTGAAGTTGCTGCA 710
 Db 651 GCCTCAGAACCGGAGCCACTTCTAGCTTGGGAGCTTGGGGCTTCCAGCCACTGTGC 456
 QY 397 TGCTTCAACCGCAGCAATTTCTGTGCTGTGAGCAGGAGAGCTTCCACCACTGTGC 456
 Db 397 TGCTTCAACCGCAGCAATTTCTGTGCTGTGAGCAGGAGAGCTTCCACCACTGTGC 456

QY 711 CCTCATCAGATTGGCCACCG---TGGGAGCATGTGCTCCACCTTGAGGCTGGACGTGT 767
 Db 457 CTTTGTGGAGATGGCCACCGGGGAGAGCGCCGTCCTCCGGCTGACCCAGGAAGAT 516
 QY 768 GGAAGATGGCCGGGGGGGGTGGCTCAGCAGCCACGCCCTCCCTTGGCAGCAGCTTCAT 827
 Db 517 AGAGATGGCAGGAGGAGGAGAGGCTATATGACCCACGAGCATGGGCTCCCTGCTGT 576
 QY 828 AGACGGGAGAGCTGTACAGGGGCTCAGCTGTCTGCTGTGGGGGAGAGGAGCCTATGCT 887
 Db 577 GGGGAGAGAGCTATCTAGGAGGGGTGACAGACCTCATGGAGAGAGCTTACATCTT 636
 QY 888 CCGAGTGGAGGTCTCTGCGCCAGCTCTGCTTCCAGCTCTGACCAAGATCTTCTTC--A 944
 Db 637 TCGCAGCCTTAGGGCAACCTCCAGTCTCCGAACAGACACACATCTCCCTGGCTCA 696
 QY 945 CGACCCCGGTTTGTGATGGCGCCGCGATCCCTGAGAACTGTGACAGAGACATGACAA 1004
 Db 697 TGACCCCAAGTTTGTCAAGATATTTTGGATTCGAGAGAGAGAGAACCCAGACAGCA 756
 QY 1005 GGTGTACTTCTTCTCTCGAGAGAGGTCCTCCGATGTGCTGCAACCATGTCTAC 1064
 Db 757 AATCTACTTCTTCTTCTGTGAGAGGGGTAGAGGGGCGCCGACACTGGAGCCTGTCT 816
 QY 1065 TGTCAAGCCGCTGGCGCCGCTGTGCGGATGATGCTGGGGGCGCAGCGGTCTGTGAA 1124
 Db 817 CGTGTCCCGTGGGCGAGATCTGCCGGAACGACGTGGCGGCGCAC--GCASCTGGTCA 875
 QY 1125 CAATGAGACACTTCTCTCAAGGCCAGCGCTGTCTGCTGCTGCC 1169
 Db 876 CAAGTGCAGCAGTCTCTCAAGGCGCGGCTGTGCTGCTGCC 920

RESULT 13

AL584324 786 bp mRNA linear EST 28-FEB-2001
 LOCUS AL584324
 DEFINITION Gallus gallus cdna clone ROS007F06, mRNA sequence.
 ACCESSION AL584324.1 GI:13163055
 VERSION EST.
 KEYWORDS chicken.
 SOURCE Gallus gallus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 1 (bases 1 to 786)
 REFERENCE 1 (bases 1 to 786)
 AUTHORS Murray, F.
 JOURNAL Stratagene Chick Embryo Lambda cDNA Library
 COMMENT Unpublished (2001)
 Contact: Frazer Murray
 Dept. Genomics and Bioinformatics
 Roslin Institute
 Roslin, Midlothian, EH25 9PS, UK
 Tel: +44 (0)131 527 4200
 Fax: +44 (0)131 440 0434
 Email: frazer.murray@bbsrc.ac.uk
 Seq primer: T3.

FEATURES

source

Location/Qualifiers

1..786

/organism="Gallus gallus"

/db_xref="taxon:9031"

/clone="ROS007F06"

/clone_1lb="Stratagene Chick Embryo Lambda cDNA Library (* 937405)"

/tissue_type="Embryo"

/dev_stage="5 days old"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Vector: pBLUESCRIPT SK; Site 1: EcoRI, Site 2: XhoI
 ; Cloned unidirectionally. Primer: Oligo dT, Uni-ZAP XR
 5' GAATTCGCGACGAG 3'; 3' adaptor sequence: 5'

BASE COUNT 166 a 242 c 218 g 155 t 5 others
 ORIGIN

Query Match 9.0%; Score 237.4; DB 9; Length 786;
 Best Local Similarity 61.7%; Pred. No. 1,4e-37;
 Matches 434; Conservative 0; Mismatches 253; Indels 16; Gaps 3;

QY 960 GATGCGCCCGGATCCCTAGAACTCTGACAGAGCAATGAGAGTGTCTTCTT 1019
 Db 1 GAATTCGCGACGAGTCTCCAGACAGAGAGAGAGAGATGATGAACTCTTCTTCTT 60
 QY 1020 CTGGAGACGCTCCCTCGCCGATGAGTGTGACATGTCATGTCAGCCGGTGG 1079
 Db 61 CCGGAGAG--TGGGAGATGCCCCGCTCACGCCCGGGGGTCTATTTCCGATGG 114
 QY 1080 CCGCTGTGCGTGAATGATGCTGGGGCCAGGGGTGTGTGAACTAATGAGACATT 1139
 Db 115 GGGCATCTCTGATGACAGAGATGAGAGACATGCTGCGTGGTGAACATGAGACCTT 174
 QY 1140 CCTCAAGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199
 Db 175 CCGTGAAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
 QY 1200 CCAGCTAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
 Db 235 TGAGCTTACAGACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1294
 QY 1260 GCTGTACAGACGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
 Db 295 CGTCTTCTCCGATCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1354
 QY 1320 AGACATCTGGAGATGCTTCAACGGGCTTGGCCACCGAGATGGCTCAGACAGTGT 1379
 Db 355 CGACATCCGATGCTGTCAATGAGGCTTGGCCACCAAGAGGGTCCCACTACAGATG 414
 QY 1380 GGGGCTTACAGGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
 Db 415 GATGCCCTTACAGGCGCAAAATGCTTACCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
 QY 1440 CGCACAGCCAGGAGGCTTGGCAGCACCAAGACATACCAATGAGTGTCTGCTGCT 1499
 Db 475 CACACCTTCATGAAATGCT-----AACCAAGACATGACCGGCAAGATCAACTT 525
 QY 1500 TCCCGGAGCCGCCCTCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559
 Db 526 CATGCGCTGTACCTCTGATGATACACGCGCTTACCGGCGGCGGCGGCGGCGGCGGCT 585
 QY 1560 TGTCAAGCCCACTGCGCCAGCAGCTACACCAAGATGCTGTGATGAGCGCTGAGAGCAGA 1619
 Db 586 CTTGGGCGCACACGCTCAACTA--CCGCTTACACCATGCTGTGTACCAAGGTGAGCGCGG 644
 QY 1620 GGATGGAGCTACATGCTTCTTCTGCGGAGCTGACTCAGG 1662
 Db 645 AGATGGCGCTATGANTGCTTTTCTGCGGCAAAATGCGGG 687

RESULT 14

AL584625 754 bp mRNA linear EST 28-FEB-2001
 LOCUS AL584625
 DEFINITION Gallus gallus cdna clone ROS012B12, mRNA sequence.
 ACCESSION AL584625
 VERSION AL584625.1 GI:13163356
 KEYWORDS chicken.
 SOURCE Gallus gallus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 1 (bases 1 to 754)
 REFERENCE 1 (bases 1 to 754)
 AUTHORS Murray, F.
 JOURNAL Stratagene Chick Embryo Lambda cDNA Library

TITLE
 Stratagene Chick Embryo Lambda cDNA Library

Db 241 GACTTACCCATCTTCGAAAGCCCTTGATGAGATCCGAGATCCGACAGAGCCCATGA 300
 QY 933 GAGTCTCTTGC---ACGACCCCGGTTTGTATGATGCGCCCGGATCCCTGAGAACTGTA 300
 Db 301 TTCCCGCTGCTCAATGAAACCAAGTTGTCAAGTCTTTTGAATCCAGAGAGTGA 989
 QY 990 CCAGACAAATGACAAAGTGTACTTCTTCTTCTCGAGACGTCCTCCGATGATG 360
 Db 361 CCTGATGACGATAAATCTTTCTTCTTCTTCTCGAGATCCGATGAGAAAGCAGACAC 1049
 QY 1050 CTCGACCATGTCTACGTCTGACCCGCTGCGCGCTGCTGCTGATGATGATGATG 420
 Db 421 AATGGGGGAGATGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1109
 QY 1110 GCGGCTGT 480
 Db 481 GCGGAGCTGT 1169
 QY 1170 GCGGCTGT 540
 Db 541 TGGAGTTGAGGGT---GACACCCACTTGTACCACTTCAGATGTCTTCTTCTTCT 1229
 QY 1230 CAAGCGCGGAGAGAGCTGTGAGGTGTACGCGCTGTGTGTGTGTGTGTGTGTGTGT 597
 Db 598 CCGAGACCGGAGAGAGCTGT 1289
 QY 1290 GCGGCTGT 657
 Db 658 GCGGCTGT 1349
 QY 1350 TGGCCACGAGATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 716
 Db 717 TGTCTACAAAGAGGAGGCTGT 1409
 QY 1410 TCGGCTGT 776
 Db 777 AAG-ACGAGATGT 1453
 AAG-ACGAGATGT 819

Search completed: October 9, 2002, 13:14:51
 Job time : 2034 secs

Thu Oct 10 09:35:02 2002

us-09-813-290-2.rag

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 12:40:53 : Search time 59 Seconds
(without alignments)
1647.281 Million cell updates/sec

Title: US-09-813-290-2

Sequence: 1 MACALAGKVPFGSPVWVK.....KMSRVVAENHRTPREVAT 875

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

747574

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4746	100.0	875	22	Novel human protein
2	4215	88.8	782	21	Human secreted pro
3	4215	88.8	782	22	Novel human protei
4	4203.5	88.6	785	21	Human secreted pro
5	4046.5	85.3	779	22	Amino acid sequenc
6	2052.5	43.2	775	19	Human semaphorin H
7	2041.5	43.0	775	20	Mouse semaphorin H
8	1942.5	40.9	777	20	Mouse semaphorin H
9	1942.5	40.9	777	22	Amino acid sequenc
10	1830	38.6	771	16	Human semaphorin I
11	1830	38.6	771	22	Amino acid sequenc

12	1810	38.1	796	19	AAV21264	Human semaphorin I
13	1747	36.8	749	22	AAV21264	Amino acid sequenc
14	1721	36.3	777	20	AAV21264	Human brain tissue
15	1721	36.3	777	21	AAV21264	Human PRO1491 (UNO
16	1721	36.3	777	22	AAV21264	Human PRO polypept
17	1721	36.3	777	22	AAV21264	Protein of the inv
18	1686	35.5	807	22	AAV21264	Amino acid sequenc
19	1652	34.8	751	22	AAV21264	Amino acid sequenc
20	1540	32.4	751	20	AAV21264	Amino acid sequenc
21	1538	32.4	751	21	AAV21264	Human semaphorin E
22	1538	32.4	751	22	AAV21264	Human semaphorin E
23	1289	27.2	477	16	AAV21264	Human semaphorin E
24	879.5	18.5	833	22	AAV21264	Human semaphorin E
25	876	18.5	833	22	AAV21264	Human semaphorin E
26	874.5	18.4	861	22	AAV21264	Human semaphorin E
27	874.5	18.4	861	22	AAV21264	Human semaphorin E
28	874.5	18.4	861	22	AAV21264	Human semaphorin E
29	867.5	18.3	862	22	AAV21264	Human semaphorin E
30	867.5	18.3	862	22	AAV21264	Human semaphorin E
31	867.5	18.3	862	22	AAV21264	Human semaphorin E
32	862	18.2	834	22	AAV21264	Human semaphorin E
33	862	18.2	834	22	AAV21264	Human semaphorin E
34	855	18.0	837	21	AAV21264	Human semaphorin E
35	855	18.0	837	22	AAV21264	Human semaphorin E
36	855	18.0	837	22	AAV21264	Human semaphorin E
37	851.5	17.9	590	22	AAV21264	Human semaphorin E
38	851.5	17.9	590	22	AAV21264	Human semaphorin E
39	851.5	17.9	590	22	AAV21264	Human semaphorin E
40	824.5	17.4	770	22	AAV21264	Human semaphorin E
41	776	16.4	777	21	AAV21264	Human semaphorin E
42	751	15.8	776	19	AAV21264	Human semaphorin E
43	750.5	15.8	761	21	AAV21264	Human semaphorin E
44	737	15.5	761	21	AAV21264	Human semaphorin E
45	737	15.5	761	21	AAV21264	Human semaphorin E

ALIGNMENTS

RESULT 1	AAV21264	standard; protein; 875 AA.
XX	AAV21264	
XX	AAV21264	
AC	AAV21264	
XX	AAV21264	
DT	07-JAN-2002	(first entry)
XX	07-JAN-2002	(first entry)
DE	Novel human protein (NHP) sequence.	
XX	Novel human protein (NHP) sequence.	
KW	NHP: novel human protein; secreted protein; semaphorin; oxytocin;	
XX	neurohypophysis; neuropeptide; gene therapy; drug screening.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PN	WO200170806-A2.	
XX	WO200170806-A2.	
PD	27-SEP-2001.	
XX	27-SEP-2001.	
PF	20-MAR-2001: 2001WO-US08834.	
XX	20-MAR-2001: 2001WO-US08834.	
PR	20-MAR-2000: 2000US-190638P.	
XX	20-MAR-2000: 2000US-190638P.	
PR	22-MAR-2000: 2000US-191188P.	
XX	22-MAR-2000: 2000US-191188P.	
PR	31-MAR-2000: 2000US-193639P.	
XX	31-MAR-2000: 2000US-193639P.	
XX	(LEXI-) LEXICON GENETICS INC.	
PA	(LEXI-) LEXICON GENETICS INC.	
PI	Walke DW, Wilgowski NL.	Turner CA, Hilbun E, Wang X, Donoho G;
PI	Walke DW, Wilgowski NL.	Turner CA, Hilbun E, Wang X, Donoho G;
PI	Walke DW, Wilgowski NL.	Turner CA, Hilbun E, Wang X, Donoho G;
XX	Scoville J;	
XX	Scoville J;	
XX	Scoville J;	
DR	WPI: 2001-611483/70.	
XX	WPI: 2001-611483/70.	
XX	N-PSDB: AAH47790.	
XX	N-PSDB: AAH47790.	
PT	New polynucleotides encoding human proteins that share structural	


```
Query Match      100.0%; Score 4746; DB 22; Length 875,
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0.000000
```

[illegible]

RESULT 2	
AAB23609	
ID	AAB23609 standard; Protein; 782 AA.
XX	
AC	AAB23609

XX	12-JAN-2001	(first entry)
XX	Human secreted protein SEQ ID NO: 18.	
XX	Human; secreted protein; cytokine; cell proliferation;	
XX	nutritional supplement; immune modulation; autoimmune disorder;	
XX	haematopoiesis regulation; tissue growth; haemostasis; inflammation	
XX	Homo sapiens.	

Key	Location/Qualifiers
Peptide	10...22
Protein	/label= signal_peptide 23...782
	/label= mature_protein
WO200409134-A1.	
24-AUG-2000.	
18-FEB-2000; 2000WO-US04340.	
19-FEB-1999; 99US-0120680.	
23-APR-1999; 99US-0298793.	
11-AUG-1999; 99US-0149639.	
23-SEP-1999; 99US-0155686.	
01-OCT-1999; 99US-0157847.	
29-NOV-1999; 99US-0167827.	
29-NOV-1999; 99US-0167823.	
15-FEB-2000; 2000US-0298793.	

(ALPH-) ALPHAGENET INC.
Valenzuela D, Yuan O, Hoffman H, Hall J, Rapteljo P;
WPI: 2000-549267/50.
N-PSDB; AAM33109.
New secreted proteins and polynucleotides encoding them, which are
derived from *Hom sapiens*, useful for therapy, diagnosis, and research,
as well as nutritional sources or supplements -
Claim 27; Page 250-253; 309pp; English.

The present sequence is the sequence of a human secreted protein. Its cDNA was isolated from an adult brain cDNA library. The proteins and coding sequences of the invention can be used in the isolation of similar genes and proteins, in the elucidation of their function *in vivo*.

Thu Oct 10 09:35:02 2002

us-09-813-290-2.1rag

CC and to treat a number of conditions. It is possible that they may have
 CC uses as nutritional supplements, as cytokine or cell proliferation
 CC factors, in immune modulation, where they may be used to treat immune and
 CC autoimmune diseases, as haematopoietic regulators (treating myeloid or
 CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
 CC have chemokine or chemotactic activity, haemostatic or thrombolytic
 CC activity, or anti-inflammatory activity.

Sequence	782 AA:	88.8%:	Score 4215:	DB 21:	Length 782:	Matches	782:	Conservative	100.0%:	Pred. No. 0:	Mismatches	0:	Indels	0:	Gaps	0:
Query Match																
Best Local Similarity																
Matches	782:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:							
94	MAPSAMAIICMLLGGILLHGGSSGSPSPVRLSLSTRDLSSANSAIFLGGPSLNLQA	60														
1	MAPSAMAIICMLLGGILLHGGSSGSPSPVRLSLSTRDLSSANSAIFLGGPSLNLQA	60														
154	MYLDEYRDRLFLGGLDALYSRLDQAMPDPREVLMPPQPGRECVKRGKRDPLTECANFV	120														
61	MYLDEYRDRLFLGGLDALYSRLDQAMPDPREVLMPPQPGRECVKRGKRDPLTECANFV	273														
214	RYLQPHNRTHLACGTAFOPTCALITVGRGHEVHLLEPGSVESGRCGRCPHEPSRPFAS	180														
121	RYLQPHNRTHLACGTAFOPTCALITVGRGHEVHLLEPGSVESGRCGRCPHEPSRPFAS	333														
274	TFIDGELYTGTLADFLGREAMIFRSGGPRPALRSDSDSLHDPFVMAARIPENSDDN	240														
181	TFIDGELYTGTLADFLGREAMIFRSGGPRPALRSDSDSLHDPFVMAARIPENSDDN	393														
334	DKVYFFSETPSPDGSNNHMYSRGRCVNDAGGQRYLVNKNSTFLKALVCSVPDGG	400														
241	DKVYFFSETPSPDGSNNHMYSRGRCVNDAGGQRYLVNKNSTFLKALVCSVPDGG	453														
394	GAETFFDLEDFVLLMKRAGKSLFYVALFSTVSADFVGFVAVCYHMADINEFNGPFAHR	360														
301	GAETFFDLEDFVLLMKRAGKSLFYVALFSTVSADFVGFVAVCYHMADINEFNGPFAHR	513														
454	DGPOHOMGGYGGKVPFRRGVCPSPKMTAQGPGRPGSTKIDPDEVLOFARAHPLMFPVRP	420														
361	DGPOHOMGGYGGKVPFRRGVCPSPKMTAQGPGRPGSTKIDPDEVLOFARAHPLMFPVRP	573														
514	RHGRPVLVKTHLAQOLHOLVYVRAEDGTVDYIFLGTDSGSLYKVIALQAGSABEDEV	480														
421	RHGRPVLVKTHLAQOLHOLVYVRAEDGTVDYIFLGTDSGSLYKVIALQAGSABEDEV	633														
574	VLEELQYFKVPTPTTEMEISVKRQMLYVGSRLGVAQLRLHOCETYGACAECCCLARDPYC	540														
481	VLEELQYFKVPTPTTEMEISVKRQMLYVGSRLGVAQLRLHOCETYGACAECCCLARDPYC	693														
634	AMDGASCHHYRPSLGRFRRODIRHGNPALQCLQSGOEEENGLVAATMVYGEHNTF	600														
541	AMDGASCHHYRPSLGRFRRODIRHGNPALQCLQSGOEEENGLVAATMVYGEHNTF	753														
694	LECLPSPQAAVRLMLORPDGDPQVKTDERVLATERGLLPRLSRFDGTYCTTLEH	660														
601	LECLPSPQAAVRLMLORPDGDPQVKTDERVLATERGLLPRLSRFDGTYCTTLEH	813														
754	GFSTQVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYDIQILGFANLP	720														
661	GFSTQVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYDIQILGFANLP	873														
814	RVDEYCRVWRCGTTECSGCFRSRSRGKQARGSNAGLELGKMKSRVAHEHNTPREVE	780														
721	RVDEYCRVWRCGTTECSGCFRSRSRGKQARGSNAGLELGKMKSRVAHEHNTPREVE															
874	AT 875															
781	AT 782															

RESULT 3
AAG5620

ID	AAG65620 standard; Protein: 782 AA	
XX		
AC	AAG65620:	
XX		
DT	07-JAN-2002 (first entry)	
XX		
DE	Novel human protein (NHP) sequence.	
XX		
KW	NHP: novel human protein; secreted protein; semaphorin; oxytocin; neurohypophyseal; neotropic; gene therapy; drug screening.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200170806-A2.	
PD	27-SEP-2001.	
XX		
PF	20-MAR-2001; 2001WO-050834.	
XX		
PR	20-MAR-2000; 2000US-190638P.	
XX		
PR	22-MAR-2000; 2000US-191188P.	
XX		
PR	31-MAR-2000; 2000US-193639P.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
PI	Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G; Scoville J;	
XX		
DR	WPI: 2001-611483/70. N-PSDB; AAA47791.	
XX		
PT	New polynucleotides encoding human proteins that share structural similarity with semaphorin proteins, protein hormones of neurohypophyseal family for drug screening, diagnosis and therapy of biological disorders	
XX		
PS	Claim 4: Page 38-40; 43pp; English.	
XX		
CC	The invention relates to novel human secreted proteins (NHP) that share structural similarity with semaphorin proteins, protein/peptide hormones of the neurohypophyseal family and oxytocin (neurophysin 1 precursor) family. The NHP nucleotide sequences are useful in drug screening techniques for treating symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body. Nucleotide constructs encoding NHP products are useful in gene therapy for modulating host cells expression. The constructs can be used to genetically engineer cells of a function as bioreactors in vivo, these genetically engineered cells are NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences also find use in molecular mutagenesis/evolution of proteins that are partially encoded by the NHP sequences, as reagents in diagnostic assays, for useful for generating antibodies, as reagents in diagnostic assays, for identifying other cellular gene products related to NHP and as reagents in assays for screening for compounds that are useful in the treatment of mental, biological or medical disorders and diseases. The present sequence represents the amino acid sequence of a NHP.	
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		

```

Db 121 RVLPNHNTHLLACGTGAFPTCALITVGHREHVLHLEPGSVESGRGRCPEHRRFPAS 180
QY 274 TFIIDGLYGLTADFLGRAMIFRSGGPPALRSDSDSLHDPFVMAAIPENSODN 333
Db 181 TFIIDGLYGLTADFLGRAMIFRSGGPPALRSDSDSLHDPFVMAAIPENSODN 333
QY 334 DKVYFEFSETVSPDGGSNHVTSRVGVNCVNDAGGQRLVNMKSTFLKARLVCSVPGP 240
Db 241 DKVYFEFSETVSPDGGSNHVTSRVGVNCVNDAGGQRLVNMKSTFLKARLVCSVPGP 393
QY 394 GAETHFQDLEDFVFLMPKAKSLLEVYALFSTVSAVFOGFAVCYHNAIDWEVNGSPFAR 453
Db 301 GAETHFQDLEDFVFLMPKAKSLLEVYALFSTVSAVFOGFAVCYHNAIDWEVNGSPFAR 453
QY 454 DGPQHOMGPRYGKVPFRPGVCSKMTAOPGRPFSTKDYDPEVLOFARHPLMFWVRP 513
Db 361 DGPQHOMGPRYGKVPFRPGVCSKMTAOPGRPFSTKDYDPEVLOFARHPLMFWVRP 513
QY 514 RHGRPVLVTKHIAQOLHQLVDRVDAEDGYDVIPLGTDSSVYLKVALAOGSAPREVP 420
Db 421 RHGRPVLVTKHIAQOLHQLVDRVDAEDGYDVIPLGTDSSVYLKVALAOGSAPREVP 573
QY 574 VLEELQVRFVPTITMEISYKRMVLYGSRGLYAOQLRHOCETTYGACAECCCLARDPYC 480
Db 481 VLEELQVRFVPTITMEISYKRMVLYGSRGLYAOQLRHOCETTYGACAECCCLARDPYC 633
QY 634 AMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLGOSQEEEAAGLVAAATMYGTENSTF 540
Db 541 AMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLGOSQEEEAAGLVAAATMYGTENSTF 693
QY 694 LECLEPKSPQAAVWMLQRPDEGPDQVKTDERVLTHERLLRRLSRDAGTYTCTLEH 600
Db 601 LECLEPKSPQAAVWMLQRPDEGPDQVKTDERVLTHERLLRRLSRDAGTYTCTLEH 753
QY 754 GFSQTVRATLVYIVASQDLNLFPEPKREPPARGCGLASTPPRAWYDIQLGPNLTP 660
Db 661 GFSQTVRATLVYIVASQDLNLFPEPKREPPARGCGLASTPPRAWYDIQLGPNLTP 813
QY 814 RVDYCERWVGKTEGCGCRRSRGKQANGKSWAGLELKKKSVHAEHNTPREVE 720
Db 721 RVDYCERWVGKTEGCGCRRSRGKQANGKSWAGLELKKKSVHAEHNTPREVE 873
QY 874 AT 875
Db 781 AT 782

```

```

PR 01-OCT-1999; 99US-0157247.
PR 29-NOV-1999; 99US-0167822.
PR 29-NOV-1999; 99US-0167823.
PR 15-FEB-2000; 2000US-0298733.
PA (ALPH-) ALPHAGENE INC.
PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
PR WPI: 2000-549267/50.
PT New secreted proteins and polynucleotides encoding them, which are
PT derived from Homo sapiens, useful for therapy, diagnosis, and research,
PS as well as nutritional sources or supplements -
PS Disclosure: Page 298-300; 309pp; English.
CC The present invention is concerned with a number of secreted proteins
CC and their coding sequences isolated from various human cDNA libraries.
CC The proteins and coding sequences can be used in the isolation of
CC similar genes and proteins, in the elucidation of their function of
CC and to treat a number of conditions. It is possible that they may have
CC uses as nutritional supplements, as cytokine or cell proliferation
CC factors, in immune modulation, where they may be used to treat immune
CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
CC have chemokine or chemotactic activity, haemostatic or thrombolytic
CC activity, or anti-inflammatory activity. No information about sequences
CC AAB23632-B23645 is given in the specification.
CC XX
Sequence 785 AA:
Query Match 88.6%; Score 4203.5; DB 21; Length 785;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 94 MAPSAMALCWLIGGLLHGGSSGSPSPVRLRLSYRDLSSANRAIFLPGDGLNLA 153
Db 1 MAPSAMALCWLIGGLLHGGSSGSPSPVRLRLSYRDLSSANRAIFLPGDGLNLA 60
QY 154 MYLDEYDRFLFLGDLALYSILRLDQAMPDPREVLMPPOGREGCRKGRDPLTCANFV 213
Db 61 MYLDEYDRFLFLGDLALYSILRLDQAMPDPREVLMPPOGREGCRKGRDPLTCANFV 120
QY 214 RVLPNHNTHLLACGTGAFPTCALITVGHREHVLHLEPGSVESGRGRCPEHRRFPAS 273
Db 121 RVLPNHNTHLLACGTGAFPTCALITVGHREHVLHLEPGSVESGRGRCPEHRRFPAS 180
QY 274 TFIIDGLYGLTADFLGRAMIFRSGGPPALRSDSDSLHDPFVMAAIPENSODN 333
Db 181 TFIIDGLYGLTADFLGRAMIFRSGGPPALRSDSDSLHDPFVMAAIPENSODN 333
QY 331 QNDKVVYFEFSETVSPDGGSNHVTSRVGVNCVNDAGGQRLVNMKSTFLKARLVCSV 300
Db 241 QNDKVVYFEFSETVSPDGGSNHVTSRVGVNCVNDAGGQRLVNMKSTFLKARLVCSV 393
QY 391 GPGAETHFQDLEDFVFLMPKAKSLLEVYALFSTVSAVFOGFAVCYHNAIDWEVNGSPF 453
Db 301 GPGAETHFQDLEDFVFLMPKAKSLLEVYALFSTVSAVFOGFAVCYHNAIDWEVNGSPF 453
QY 451 AHRDGPQHOMGPRYGKVPFRPGVCSKMTAOPGRPFSTKDYDPEVLOFARHPLMFWVRP 513
Db 361 AHRDGPQHOMGPRYGKVPFRPGVCSKMTAOPGRPFSTKDYDPEVLOFARHPLMFWVRP 513
QY 511 VPRRGRPVLVTKHIAQOLHQLVDRVDAEDGYDVIPLGTDSSVYLKVALAOGSAPREVP 420
Db 421 VPRRGRPVLVTKHIAQOLHQLVDRVDAEDGYDVIPLGTDSSVYLKVALAOGSAPREVP 573
QY 571 BEVLEELQVRFVPTITMEISYKRMVLYGSRGLYAOQLRHOCETTYGACAECCCLARD 630
Db 481 BEVLEELQVRFVPTITMEISYKRMVLYGSRGLYAOQLRHOCETTYGACAECCCLARD 630
QY 631 PYCAMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLGOSQEEEAAGLVAAATMYGTENH 690

```

|||||
 541 PYAMGASCTHYRRSLGRFRFRDINRGNPALOCGQSEEAVALVATVYSTEHN 600
 QY STELECTPKSPQAAVWILORPGDEGPDQVKTDERVLTHTENGILLFRRLSPFDAGTYCTT 750
 691 STELECTPKSPQAAVWILORPGDEGPDQVKTDERVLTHTERKLLFRRLSRDAGTYCTT 660
 Db STELECTPKSPQAAVWILORPGDEGPDQVKTDERVLTHTERKLLFRRLSRDAGTYCTT 660
 601 STELECTPKSPQAAVWILORPGDEGPDQVKTDERVLTHTERKLLFRRLSRDAGTYCTT 660
 QY LEHGFSGVWVRLALVIVASQDNLFPPEPPPEPPARGLASTPPKAWYDIIOLIGFA 810
 751 LEHGFSGVWVRLALVIVASQDNLFPPEPPPEPPARGLASTPPKAWYDIIOLIGFA 720
 Db LEHGFSGVWVRLALVIVASQDNLFPPEPPPEPPARGLASTPPKAWYDIIOLIGFA 810
 QY 811 NLPVDECEVWVCGTTECSGCFRFRSRSRQKARSKMAGLELGGKMKSRVHAENRTPR 870
 721 NLPVDECEVWVCGTTECSGCFRFRSRSRQKARSKMAGLELGGKMKSRVHAENRTPR 780
 Db 871 EYEAT 875
 QY 781 EYEAT 785
 Db 781 EYEAT 785

RESULT 5
 AAB84219
 ID AAB84219 standard; protein; 779 AA.
 XX AAB84219;
 AC 06-AUG-2001 (first entry)
 XX
 DT
 XX
 DE Amino acid sequence of a semaphorin polypeptide designated ZSMF-16.
 XX
 XX Human: semaphorin; ZSMF-16; neurite growth; neurite outgrowth;
 XX T lymphocyte suppressor; cancer; stroke; brain damage; paralysis;
 XX spinal injury; neurodegenerative disease; amyotrophic lateral sclerosis;
 XX Alzheimer's disease; Huntington's disease; Parkinson's disease;
 XX peripheral neuropathy; demyelinating disease; multiple sclerosis;
 XX immunosuppression; autoimmune disease; insulin dependent diabetes;
 XX rheumatoid arthritis.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..22 "signal peptide"
 FT /note="signal peptide"
 FT 23..779
 FT Protein /note="mature protein"
 FT 62
 FT Modified-site /note="N-glycosylation site"
 FT 76..500
 FT Domain /note="semaphorin"
 FT 124
 FT Modified-site /note="N-glycosylation site"
 FT 593..654
 FT Domain /note="Ig-like domain"
 FT 594
 FT Modified-site /note="N-glycosylation site"
 FT 594
 FT
 XX WO200140278-A2.
 XX
 XX 07-JUN-2001.
 XX
 XX 06-DEC-2000; 2000MO-US33116.
 XX
 XX 06-DEC-1999; 99US-0455560.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Holloway JL, Foley KP;
 XX
 XX WPI: 2001-374784/39.
 XX
 XX N-PSDB; AAF90250.
 XX
 XX Novel human semaphorin polypeptide, ZSMF-16, useful for treating
 XX peripheral neuropathies Alzheimer's and Huntington's disease and
 PT

polynucleotide encoding ZSMF-16 useful for detecting genetic
 abnormality and cancer -
 Claim 13; Page 121-123; 124pp; English.
 The present sequence represents a semaphorin polypeptide, designated
 ZSMF-16. ZSMF-16 is a neurite growth and development modulator. It also
 enhances spinal cord and sensory neurite outgrowth and patterning, and
 is involved in the activation and regulation of T lymphocytes suppressor.
 ZSMF-16 is useful for detecting a genetic abnormality or cancer. ZSMF-16
 polynucleotide probes can be used to detect 3p21 loss, trisomy,
 duplication or translocation associated with mammary tumor tissue,
 breast tumor, liver, small intestine, bone cancers, etc.. ZSMF-16 can
 be used to modulate neurite growth and development and demarcate nervous
 system structures. ZSMF-16 are also useful for regenerating and directing
 neurite outgrowths following strokes, brain damage caused by head
 injuries, paralysis caused by spinal injuries, and for treating
 neurodegenerative diseases such as amyotrophic lateral sclerosis,
 Alzheimer's disease, Huntington's disease, Parkinson's disease, multiple
 sclerosis. ZSMF-16 also acts as a mediator of immunosuppression,
 and thus useful for diagnosing and treating autoimmune diseases such as
 insulin dependent diabetes, rheumatoid arthritis, and multiple
 sclerosis. It can also be used as an anti-inflammatory for inhibition
 of antigen in humoral and cellular immunity and for immunosuppression
 in graft and organ transplants.

Sequence 779 AA:
 Query Match 85.3%; Score 4046.5; DB 22; Length 779;
 Best Local Similarity 95.0%; Pred. No. 0; Indels 39; Gaps 2;
 Matches 760; Conservative 0; Mismatches -----ILS 135

94 MAPSMAICWLLGLLHGSSGSPGSPVPRRLSYRD-----ILS 135
 1 MAPSMAICWLLGLLHGSSGSPGSPVPRRLSYRDAMWRKPSMTMMETFSRYLS 60
 136 ANRSALIFLGPGSLNLOMYLDEYRDLFLGDLALSLRLDQAWDPREVLMPOQGR 195
 61 ANRSALIFLGPGSLNLOMYLDEYRDLFLGDLALSLRLDQAWDPREVLMPOQGR 110
 196 EECVRRGRDPLTECANFVRLDPHNRTHLACGTCGAFOPTCALITVGHREHYHLEPGS 159
 111 -----TECANFVRLDPHNRTHLACGTCGAFOPTCALITVGHREHYHLEPGS 159
 256 VESGRGCPHEPSRPFASFIDELTYGLTADFLGEMIFRSRGGPRALRSDSOLH 315
 160 VESGRGCPHEPSRPFASFIDELTYGLTADFLGEMIFRSRGGPRALRSDSOLH 219
 316 DREFVMAARIPENSDDNDKYFFEFSETPSPDGSNHYTVSRVGVYCNADGGRVLYN 279
 220 DREFVMAARIPENSDDNDKYFFEFSETPSPDGSNHYTVSRVGVYCNADGGRVLYN 279
 376 KWSTFLKARLYCSVPDGAETHFDQLEDVFLMPKAGKSELYVALFSTYSAVFOGFANC 339
 280 KWSTFLKARLYCSVPDGAETHFDQLEDVFLMPKAGKSELYVALFSTYSAVFOGFANC 339
 436 VYHMADIMEVFNCPFAHRDGPQHOMGPYGGKYPFRPGVCSKMTAOPGRFGSTKDPD 495
 340 VYHMADIMEVFNCPFAHRDGPQHOMGPYGGKYPFRPGVCSKMTAOPGRFGSTKDPD 399
 496 EYLOFARAHPLMEVPVRGRGRPVLVKTHLAQDLHIVDRKADGDTYDVFILGTDSGS 555
 400 EYLOFARAHPLMEVPVRGRGRPVLVKTHLAQDLHIVDRKADGDTYDVFILGTDSGS 459
 556 VLKVALDAGGSAPEEVEVLELOVFKVPPITMEISYVRKMLVYGSRLGVNOJRLHQC 615
 460 VLKVALDAGGSAPEEVEVLELOVFKVPPITMEISYVRKMLVYGSRLGVNOJRLHQC 519
 616 EYGTACACCCCLARDPYCAMDASCTHYRPSLGRFRFRDINRGNPALOCGQSEEA 579
 520 EYGTACACCCCLARDPYCAMDASCTHYRPSLGRFRFRDINRGNPALOCGQSEEA 579

QY 676 VGLVAATVYGTENHSTFLECKRSPQAAVWMLQRPDEGPDYKTDERVLHMERGLLF 735
 Db 580 VGLVAATVYGTENHSTFLECKRSPQAAVWMLQRPDEGPDYKTDERVLHMERGLLF 735
 QY 736 RLRSFAGTYTCTTLEHSPQTVVRLVAVASQDNLFPPEPKPEEPARGGLASTP 795
 Db 640 RLRSFAGTYTCTTLEHSPQTVVRLVAVASQDNLFPPEPKPEEPARGGLASTP 795
 QY 796 PKAWYDIQILIGFANLPVDEYCEWVRCGTTCSCGFSRSRSGKQARKSMAGLELCK 855
 Db 700 PKAWYDIQILIGFANLPVDEYCEWVRCGTTCSCGFSRSRSGKQARKSMAGLELCK 855
 QY 855 KKKSVNAEHNTPREVEAT 875
 Db 760 KKKSVNAEHNTPREVEAT 875

RESULT 6
AAW63748

ID AAW63748 standard; Protein; 775 AA.

AAW63748:

01-OCT-1998 (first entry)

Human semaphorin.

Human; semaphorin; diagnosis; nerve disease; nerve growth inhibitor;
 neurological disease; atopic skin inflammation; autoimmune disease;
 pain.

Homo sapiens.

MO9822504-A1.

28-MAY-1998.

12-NOV-1997; 97WO-JP04111.

15-NOV-1996; 96JP-0321068.

(SDMU) SUMITOMO PHARM CO LTD.

Furuyama T, Inagaki S.

WPI; 1998-312416/27.

N-PSDB; AAV35367.

Gene encoding new semaphorin nerve growth inhibitor - useful in
 diagnosis, treatment and study of neurological diseases
 Claim 1; Page 33-37; 49pp; Japanese.

The present sequence represents human semaphorin, a nerve growth
 inhibitor. The semaphorin protein, and gene encoding the protein,
 and their derivatives, are used in the diagnosis, treatment and
 autoimmune diseases and pain.

Sequence 775 AA:

Query Match 43.2%; Score 2052.5; DB 19; Length 775;
 Best Local Similarity 50.7%; Pred. No. 8.9e-177;
 Matches 399; Conservative 127; Mismatches 234; Indels 27; Gaps 10;

QY 94 MASAAMICWILGILLHCGSSGSPGSPVRLRLSVRDLLSARSAIFLPGQSLNIA 153
 Db 1 MASAAMICWILGILLHCGSSGSPGSPVRLRLSVRDLLSARSAIFLPGQSLNIA 153
 QY 154 MYLDEYRDRLFLGGLDALYSRLDQAMPDPREVLPMPQPGQRECVAKGRDPLTECAN 213
 Db 61 MYLDEYRDRLFLGGLDALYSRLDQAMPDPREVLPMPQPGQRECVAKGRDPLTECAN 213

QY 214 RVLQPHNTHLLACGTAFQPTCALITYGHRGHVLT-HLEPGSVSGRQCPHEPSRFA 272
 Db 120 RVLQPHNTHLLACGTAFQPTCALITYGHRGHVLT-HLEPGSVSGRQCPHEPSRFA 272
 QY 273 STFDGLYGLYGLADFLGREAMIFRSRGGPRPALRSD-IDQSLHDPREPVMAARIPNSDQ 331
 Db 180 STFDGLYGLYGLADFLGREAMIFRSRGGPRPALRSD-IDQSLHDPREPVMAARIPNSDQ 331
 QY 332 DNDKVFEESEIYPPSDGSGNHVTVSRVGVNDAGQARLVKMSSTFLKARLVCSVG 391
 Db 240 DNDKVFEESEIYPPSDGSGNHVTVSRVGVNDAGQARLVKMSSTFLKARLVCSVG 391
 QY 392 PGAEETHEDQLDVFLLMPKAKSLEYALFSPYSAVFGQFVAVVHMADIWEVNGRPA 451
 Db 299 PGAEETHEDQLDVFLLMPKAKSLEYALFSPYSAVFGQFVAVVHMADIWEVNGRPA 451
 QY 452 HRDGFQHQMGPGYGGKVPFPPRGVCPKMTAOPGRPGSTKDYPPDEVLOFARHPLMEWPV 511
 Db 359 HRDGFQHQMGPGYGGKVPFPPRGVCPKMTAOPGRPGSTKDYPPDEVLOFARHPLMEWPV 511
 QY 512 RPRHGRPVLVKTHLAQQLQIYVDEVEAEDGTVDYIFIGTSGSVLYALQAGSAEPE 571
 Db 416 RPRHGRPVLVKTHLAQQLQIYVDEVEAEDGTVDYIFIGTSGSVLYALQAGSAEPE 571
 QY 572 EYVLELQVFKVPTPTTEMTISVKRMLYVGRVAGVQALRHOCETVGTACACCLARP 631
 Db 476 EYVLELQVFKVPTPTTEMTISVKRMLYVGRVAGVQALRHOCETVGTACACCLARP 631
 QY 632 YCAMDASCCTHYRP--SLGRPRPRODIRHGNPALQCLQSOEEAVGLVATVYGTEN 689
 Db 536 YCAMDASCCTHYRP--SLGRPRPRODIRHGNPALQCLQSOEEAVGLVATVYGTEN 689
 QY 690 NSTFLECLPKRSPQAAVWMLQRPDEGPDYKTDERVLHMERGLLFRLSPFDAGTYTCT 749
 Db 596 NSTFLECLPKRSPQAAVWMLQRPDEGPDYKTDERVLHMERGLLFRLSPFDAGTYTCT 749
 QY 750 TLEHGFQTVVRLVAVASQDNLFPPEPKPEEPARGGLASTPARGGLASTPARGGLASTP 803
 Db 656 TLEHGFQTVVRLVAVASQDNLFPPEPKPEEPARGGLASTPARGGLASTPARGGLASTP 803
 QY 804 LQILGFANLPVDEYCEWVRCGTTCSCGFSRSRSGKQARKSMAGLELCKMKSRVHA 863
 Db 714 LQILGFANLPVDEYCEWVRCGTTCSCGFSRSRSGKQARKSMAGLELCKMKSRVHA 863
 QY 864 EHNKTRP 870
 Db 764 EHNKTRP 870

RESULT 7

AAV43090

ID AAV43090 standard; Protein; 775 AA.

AAV43090:

05-JAN-2000 (first entry)

Mouse semaphorin H (Sema H) amino acid sequence.
 Semaphorin H; Sema H; Sema HV; collapsing; bone structure formation;
 metastasis; cancer; antibody; drug screen.

Mus sp.

Key Location/Qualifiers

Misc-difference 615

/note="Encoded by GYA"

23-SEP-1999.

12-MAR-1999; 99WO-IB00495.

CC metastasis. The invention uses the mouse *Sema H* gene in the diagnosis of
 CC metastatic cancer. Semaphorin polynucleotides and polypeptides are used
 CC in the methods of the invention, the polypeptides are useful for
 CC determining the metastatic potential of cells, by detecting their
 CC expression in biological samples. Antibodies specific for *Sema H* are
 CC also useful therapeutically in inhibiting *Sema H* polypeptide activity and
 CC therefore metastasis, and for purifying the polypeptides. Metastasis may
 CC also be inhibited by inhibiting the biological activity of the
 CC polypeptide using e.g. a small molecule inhibitor or a semaphorin-H
 CC ligand (or fragment). The polynucleotides can also be used to inhibit
 CC prevent metastasis of cancer cells using known antisense technology e.g. to
 CC quantify *Sema H* mRNA levels in cells. They can be used to detect and
 CC multimeric proteins, antibodies or antisense oligonucleotides can be
 CC included in pharmaceutical compositions. The polynucleotides can be
 CC to isolate similar sequences from other species and to produce mammalian
 CC cell lines and tumours with known metastatic potential, useful in
 CC anti-metastatic drug screening.

Sequence 777 AA:

Query Match

Best Local Similarity 40.98; Score 1942.5; DB 20; Length 777;
 Matches 387; Conservative 125; Mismatches 248; Indels 29; Gaps 12;

94 MAPSAMALCMLGGLLHGSSGSPSPVRLRLSYRDLNSANSAFLGPOSLNQA 153
 1 MAPAGHILTLMLGHLLEMTWPGHSANPSYARLPISHKELEFELNGLOTFKAPLGFIDLHT 60
 154 MYDEYDRFLGLGDLALSLRLDAMPDPREVLPPOGORECVCVKGRDPLTECANEF 213
 61 MLDEYQERLFLGGRDLYVSLNLERVSDGYREIYWPSTAVKVECIKMGD-ANECANYI 119
 214 RVLQPHNTHLLACGTGAFOPICALITYGHRGEHVL-HLEPGVESGRGRCHEPSPRA 272
 120 RVLHYNTHLLCATGAFPHCAFIVGHNSEDEPLHLSHRSERGRRCFEDNSFV 179
 273 STFDIELYGLTADPLGREAMIFRSGCRPRALRSD-SQSLHDPREVMARIPENDQ 331
 180 STVGNELFLGGLSDWGRSALFIRSMGRLGHTHEDERLKEPKFVGSMYLPDNDR 239
 332 DNDKVFSEFSETPSPDGSNNHTVSRYGRCVNDAGORVLVNMSTFKARLYCSPG 391
 240 DDKMTFFETFEKALEENNA-HTILHPSGRICLVNDMGORILVNMSTFKARLYCSPG 391
 392 PGCAETHFDQLEDFVLEMPKAGKSLVYALFSTVSAFQGVAVYHMADWFEVNGPRA 451
 299 MNGIDTYDEDELDVFLPTRPKNPVIFGLFNTSNIFRGHNAVCIYHMSIRAFNGPRA 358
 452 HRDGHQMGFYGGKVRPRRPGVCPSCMTAOPRRPGSKDVRDEVLOFARAHPLMPRV 511
 359 HKEBPEYHMSLYEKKVYPRPRSCASKVN---GKKGITQRLRDALIFRAPHNPLMQFI 415
 512 RPRGRVULKTHLAQOLQIYVDRVEADGTYDIFIGTDSGSV-LKVIALQAGSAPR 570
 416 KPVHKRPILVKTIDGKYNLAQVDRVEADQYDLFGTDTGTGLVLTITNOETEMM 475
 571 EEVULEELOVFPVPTITEMEISYKRMUYGSLGVAOLRLHOCSEYTGACAECLLARD 630
 476 EEVULEELOVFPVPTITEMEISYKRMUYGSLGVAOLRLHOCSEYTGACAECLLARD 630
 631 RYCDAGASCTHNPSSG--KRRFRDRIHGNAPALOCGQGOSEBANGVLAVATWVTE 688
 536 RYCDAGASCTHNPSSG--KRRFRDRIHGNAPALOCGQGOSEBANGVLAVATWVTE 688
 689 HNSTFELCERKPOAVRMLLORPGDEGPDQYKTERVILHTERGLRRLSRFAGIYTC 595
 596 SNSTLLECPRLSLQAKVIVFLQKGRDVREKVEKTDKRVAKMDLGLLELRVAKSDAGIYTC 655
 749 TTEHEFSTVRLALVIVASQDLNLFRRPRKPE-----PARAGLSTPRKMYKD 802
 656 QVVEHNFVHTAKITLLEVVEHKEVGMFNKDHBERHNKMRCPRLSMGSGTAP--WYKE 713

OY 803 ILQILGFAN-LPRVDEYCEYWCGRTEGSGFRSRGKQARQKSNAGLELGRKMSRV 861
 DB 714 FLQILGSGKFGQVEYERCEWCTD-----KRRKTKSPSKMYANPQEKRLRSK- 764
 OY 862 HAEHNRTPR 870
 DB 765 -AEHFRPLR 772

RESULT 9

AAG62730

ID AAG62730 standard; peptide; 777 AA.

AC AAG62730;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of mouse semaphorin *Sema3E*.

KW Hanatoxin; tarantula; toxin; voltage-gated potassium channel;

KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

OS semaphorin; dorsal root ganglion repulsion; growth cone collapse.

Mus sp.

WO200138491-A2.

31-MAY-2001.

07-NOV-2000; 2000WO-US41943.

PR 08-NOV-1999; 99US-0164056.

PA (GENO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

DR WPI; 2001-451494/48.

PT Polypeptide sequences that encompass the hanatoxin-like sequences of
 PT semaphorins, useful as a drugs to treat any condition or disease that
 PT is characterized by abnormal calcium channel function
 PS Claim 6; Page 13; 29pp; English.

CC The present sequence represents a semaphorin. The specification
 CC describes hanatoxin-like sequences (HTLS) found in the semaphorin
 CC domain of mammalian secreted semaphorins. Hanatoxin is a tarantula
 CC toxin that selectively blocks some voltage-gated potassium and calcium
 CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
 CC and growth cone collapse activities associated with semaphorins
 CC Polypeptides containing HTLS can be used to modulate the activity of
 CC calcium channels. The peptides can also be used as an antigen to
 CC generate antibodies that can then be used to modulate the activity of
 CC the peptides or antibodies by inactivating naturally occurring channel ligands.
 CC or disease that is characterized by abnormal calcium channel function.

Sequence 777 AA:

Query Match

Best Local Similarity 40.98; Score 1942.5; DB 22; Length 777;
 Matches 387; Conservative 125; Mismatches 248; Indels 29; Gaps 12;

94 MAPSAMALCMLGGLLHGSSGSPSPVRLRLSYRDLNSANSAFLGPOSLNQA 153
 1 MAPAGHILTLMLGHLLEMTWPGHSANPSYARLPISHKELEFELNGLOTFKAPLGFIDLHT 60
 154 MYDEYDRFLGLGDLALSLRLDAMPDPREVLPPOGORECVCVKGRDPLTECANEF 213
 61 MLDEYQERLFLGGRDLYVSLNLERVSDGYREIYWPSTAVKVECIKMGD-ANECANYI 119
 214 RVLQPHNTHLLACGTGAFOPICALITYGHRGEHVL-HLEPGVESGRGRCHEPSPRA 272

Db	8	VC-LFENGUILLTARANYONGKNNVPRUKSTYKEMIESNNVITFNGLANSSSYHFFLLDER	66
QY	161	DRLEFLGDLALYSLRIDQWAMPREVLMPQPGORECVCYRGDRPULTECANFVRVLOPHN	22
Db	67	SRLYAGAKDHIFSPDLVNI-KDFOKIWPVPVSYSTRDECKMAGDILKECANIKYKLAVN	125
QY	221	RHHLACGGGAOPCALITVGHKGE-HVHLTPGSGVEGGRGCRHEPSRPFASTEJIGE	274
Db	126	QHHLACGTAHFHICTYIEIGHHPEDONFKLENSHFENGKRSYDPKLLTASLLDGE	185
QY	280	LYTGLTADLELAREMIRSGRPRALRPSNC-POG	

106 LYSSTADDFMGKRFALFRTLLGHHNPLRIEIQHDSRMLANDPKFISAHLLISEEDNEDDKVVF 339
107 339 FFSSTVSPGPG -GSNHVYVSRRGVCVNDAGGQRYLVAKMSTPLFKALVYCSVPGGAET 245
108 246 FFRREN -AIDGEGSGATNARIIGQICKNDPEGHRSILVKKMTTLFLAKLILCSVPGNGIPT 397
109 398 HFOQLDQVLLMPKAGKSLLEVLALESTVSAVFOGFANCVYIMHMLDIENVNGFPAHDDQ 457
110 457 304 HPELDQVFLMFKPKRPNVUVFTSSNTEFGSAVCSMSDVARVFLGPIYHRDPN 363
111 363 458 HQMGRRGKVPFRPGVCPSEKMTQDQRPKPGSTKDYRDEVLQFARAHPLMPVAPRNGR 517
112 517 364 YQWVYQGRVPRPRPGCTKTEG---GFDSTKDLRDPDVIIFARSHAPNVPFPMNRR 419
113 419 518 PVLVTHLQOOLHIVDRAEDGTVDVILGTSBGSLYKIALQAGSAPEEVLVEE 578
114 578 420 PIVKTDVNNQFQIVYDRADEDQYDVMFQIDTVGLVYLVKVSIPKEWYDLEELVEE 479
115 479 480 MYVRETPALISAMELSTKQOOLYSTAGVADQLPLHRCDIYGKACABCCCLARDPYCAMDG 637
116 637 638 ASCTHVRPSLGRKRRPRDRIHGNALQCL-----GQSQEAVAGLYAATWYGTGH 539
117 539 540 SACSRRFPT-ARRRRRQDIRGDELHSCSLNHDHNGHSRPER-----IIVGYEN 590
118 590 690 NSTPTECPKPSQAQVAKMLQRPDESDPDQKIDERYLTENGKLLFRRLSRPDATYGTCT 749
119 749 591 SSTPLECPKSORALYVQFQRRNEERKEIRVDHIIITLDGGLLRSLLOOKDSQNYLCH 650
120 650 750 TLEHGSQTVRLAVLIVASOLDNLFPRPKDEEPRPARGGLASTPP--KMYQDIQLI 807
121 807 651 AVEGFEIOTLLKLVLEIYDTEHLEELLHDDGDGSKTKENSMTSPQAKWYIDPFQOLI 710
122 710 808 GFANLPRVDECEWMCRGTECSGCFGRSRSRGQARQS--MAGLELGKKMSRYNAHEH 865
123 865 711 NHRPLNTMDFFCQVAKDKRKO-----RRQRPCHTSPNSKKMKHLDENKKGNRRTHFF 764
124 764 866 NRRPRREV 872
125 872 765 ERAPRSV 771

Human: semaphorin III wild type protein fragment 1.
 frameshift mutation; age-related disease; beta-APP; diagnosis: cancer.
 Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 oligosaccharide maltose; multiple sclerosis; alcoholic liver disease;
 oligodactyl B; apolipoprotein E; APP; microtubule associated protein; Tau;
 neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 filial fibrillary acidic protein; GFAP; p53; semaphorin III; HspF-1;

bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene. HMGP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX Homo sapiens.
 XX M09845322-A2.
 XX 15-OCT-1998.
 XX 02-APR-1998; 98MO-1B00705.
 XX 10-APR-1997; 97US-0043163.
 XX (UYUT-) RIJKSUNIV UTRECHT.
 XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 XX (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX Burbach JPH, Grosveld FG, Van Leeuwen FW.
 XX WPI; 1998-609901/51.
 XX N-PSDB: AAX575767.
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 XX corresponding protein mutations - used to diagnose cancer and
 XX neurological diseases, particularly Alzheimer's disease, and also
 XX for treatment and prevention with specific ribozymes or wild-type
 XX RNA
 XX Disclosure: Figure 16; 258pp; English.
 XX This invention describes a novel method for the diagnosis of a disease
 XX caused by, or associated with, an RNA molecule that has a frameshift
 XX mutation. The method is used to diagnose age-related diseases, especially
 XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 XX and many others listed) or susceptibility to these disorders. The method
 XX allows a definitive diagnosis of Alzheimer's disease in living patients,
 XX at an early stage. It is based on the observation that disease may be
 XX caused by mutations in RNA rather than DNA. The invention describes the
 XX use of neuronal system RNA molecules, specifically proteins including
 XX beta-amyloid precursor protein (beta-APP), the microtubule associated
 XX protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 XX neurofilament-E, presentin I, presentin II, glial fibrillary acidic
 XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 XX protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX Sequence 796 AA:
 XX 38.1%; Score 1810; DB 19; Length 796;
 XX Query Match 45.4%; Pred. NO. 9.2e155; Indels 58; Gaps 14;
 XX Best Local Similarity 139; Mismatches 244;
 XX Matches 366; Conservative 139;
 OY 101 ICMILGILLHGSSGSPSPVRLRLSYRDLSSANSRAIFLGFQSGLSLQANAYLDEYR 160
 OY 13 VC-LFMGVLLTARANYQNGKNNVRLKLSYKEMLENNVITFNGLSSSYHTFLADEER 71
 OY 161 DRFLPGDALYSRLDOAMPREYVLPPOGQRECYRKGRDPLTECANFVRLQPHN 220
 OY 72 SRLIYGADHIFESFDLVNI-KDFOKIWPVSYTRDECKKAGDLKECANFKIKYLKIN 270
 OY 221 RTHLLACGTGAFOPTICALITYGHRGE-HVLLHLEPGSVESGRKCPHEPSRPFASFIDE 279
 OY 131 QTHLYNCGTGAHPICTYIEIGHNEHDNIFKLENSHENGSGSPYDKILLTSLIDDE 190
 OY 280 LYTGLADELGRFAMTFSGGPRALRSDS-DOSLHDPFRVMAARIENSDODNKKYF 338
 OY 191 LYSTADAEFGGRDFAIFRLGHHHFKTEQDHSRWLNPKFISAHLSIESNPDEDDYVF 250
 OY 339 FFSFTVSPDG-GSNHVTYSRVGVCVNDAGORVLVKKWSPFLKARLVCSVPGGAET 397

Db 251 FREN--AIDGESHKATHARIGQICKNDPFGHSLVYKKTTFILKARLICSVDGPNIGDT 308
 OY 398 HFDOLEDEYFLMRKAGKSLVYALFSTVSAVQGFACVYHMAIDWEVNGFPAHRDQ 457
 OY 309 HFDELODFLNMFRDPAKPNVYGVFTTSSNIFGSAVCMSYSDVRVFLGYPVHROGN 368
 OY 458 HQMGPGYGGKVPFPPRGVCPKMTAQRGPRGSKDYDEVLQFARAPLMEVPR 513
 OY 369 YQWVPYQGRKVPFPRGTCPSKTEG---GFDSTRKDLDDYIFPARSHPMYMPVPMNR 424
 OY 514 -----RGRPVLYKTHLAQQLHOIYVDRVAREDGTFYDVIPLGDSGVL 557
 OY 425 PIVIKTDVNTQFTQIYMMNRPPIVITKTDVNTQFTQIYVDRVAREDGTFYDVIPLGDSGVL 484
 OY 558 KYIALQAGSAEPEYVLELOEKFVPTITEMELSVKROMLYGSGALGVAOLRHOCET 617
 OY 485 KVSIPKEMTWYDLEEVLEEMTVFEPTAISAMELSTKOOLYIGSTAGVAQLPIRHCDI 544
 OY 618 YGTACACCCCLARDPYCAMDGASCTHYRPSLGRKRRPRDRIHGNALQCT-----GQ 669
 OY 545 YKACACCCCLARDPYCAMDGASCSRYEP-ARRTRRODIRNGDPLTCHSDLHNDHGH 603
 OY 670 SOEEAVGLVAATVYGFENSTFLECLPSPQAAVRLQRPDGEPPDYKTERVLT 729
 OY 604 SPEER-----IITGVENSTFLECPSPKORALYVQFQRNEERKEIVDDHIRT 655
 OY 730 EGGLEFRLSRPDAGTYCTTLENGFSOTVRLALVYVASQDLNLFPEPPRPEPPARG 789
 OY 656 DQGLLRSLQKDSGNYLCHAVEGEIOTLKVLEYIDREHLEELHKDDGSKTKE 715
 OY 790 GLASTPR--KAWYKDIOLIGFANLPRVDEYCEYVCRGTECGCFRSRSGKQARGS 847
 OY 716 MSNSMTPSOKWYRDMQLINPNTIMDEFEQYWKDRQ-----RRORGHHPGNS 769
 OY 848 --WAGLELGKMKSRVHAENHPRREV 872
 OY 770 NKWKHLQENKGRNRRTHERAPRSV 796
 Db 770 NKWKHLQENKGRNRRTHERAPRSV 796
 RESULT 13
 AAG62727 standard; peptide: 749 AA.
 ID AAG62727
 AC AAG62727
 XX 17-SEP-2001 (first entry)
 DE Amino acid sequence of human semaphorin Sema3B.
 XX KW Hanatoxin; tarantula; toxin; voltage-gated potassium channel;
 KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;
 KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.
 XX Homo sapiens.
 OS WO200138491-A2.
 PD 31-MAY-2001.
 PF 07-NOV-2000; 2000WO-US41943.
 PR 08-NOV-1999; 99US-0164056.
 PA (GENO) GEN HOSPITAL CORP.
 PI Behar O, Woolf CJ;
 PT WPI; 2001-451494/48.
 XX polypeptide sequences that encompass the hanatoxin-like sequences of
 XX semaphorins, useful as a drugs to treat any condition or disease that
 XX is characterized by abnormal calcium channel function -

PS Claim 6; Page 12; 29pp; English.

CC The present sequence represents a semaphorin. The specification
CC describes hanatoxin-like sequences (HTLS) found in the semaphorin
CC domain of mammalian secreted semaphorins. Hanatoxin is a tatanula
CC toxin that selectively blocks some voltage-gated potassium and calcium
CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
CC and growth cone collapse activities associated with semaphorins.
CC Polypeptides containing HTLS can be used to modulate the activity of
CC calcium channels. The peptides can also be used as an antigen to
CC generate antibodies that can then be used to modulate the activity of
CC calcium channels by inactivating naturally occurring channel ligands.
CC The peptides or antibodies can be used as drugs to treat any condition
CC or disease that is characterized by abnormal calcium channel function.
XX

SO Sequence 749 AA;

Query Match
Best Local Similarity 36.8%; Score 1747; DB 22; Length 749;
Matches 367; Conservative 107; Mismatches 239; Indels 66; Gaps 17;

DB 83 GRORCPQFSPMSAMALICWLLGGLLHGGSSGSPSPVRLRLRYRD-----L 134
DB 2 GRAGAAVIFGLA-LIMAV-----GL-----GSAASP-----PRLRLSFOELQAMHGLQTF 47
OY 135 SANRSAIFIGPGSGLQAMYLDEYRDLTGLDALYSLRLDQAMPREVLPPOGQ 194
DB 48 SLERTCCY-----QALLVDEBERGLFVGAEHNVASLMDNISRAKRLAMPAREV 98
OY 195 REECYRKGRDPLTECANFVRLQPHNTHLACGAFOPCALITVGHNGH-VLHLEP 253
DB 99 RECMWAKKIDGTCECMNVKLLHAYNTHLACGAFHPTCAFEVGHAEPEVRLDP 158
OY 254 GSVEGRCRCHPSPSPASTFIDELIYGLDDELREAMIFRSGGPRALRSD-SDOS 312
DB 159 GRIDGKSPYDRHRAASVLYGEEELYSVADLMGRDFTIRSLGQPSLTERHDSR 218
OY 313 LHPDPRVMAARIPENDODNDKVFYEFSETPSPDGSNHTYSGVAVCNADGGGRV 372
DB 219 WLNPKRVKVFYLPESBNPDCKIYFFEFRETAVEAPALGRISVSRVQICNDVGGORS 278
OY 373 LVNKNSTFLKARLYCSVPFGAETHFDLEPLMLRACKSLEVALFETSVAPFGGF 432
DB 279 LVNKNSTFLKARLYCSVPFGEG-DTHFDLQDVLFLSSRDHRTPLLYAVFTSSIRG 337
OY 433 ANCVYHMDIWEVFNPAHNRDGRPOHNGPYGKVPFPPGVCSSKMTAQDGRFGSTKD 492
DB 338 ANCVYSMDVRAFLGPRFANHEGPHQVSYOGHVPYPRGMCPSKTFG-----TFSTKD 393
OY 493 YPDEVLOFARAHPLMFVPRHGRPVLYKTHLAQQLHOIVDVRVEAEDGTYYVIFLGD 352
DB 394 PPDDVIOFARAHPLMYNSVLPFGGRPLFLQVANYTTOIAADRVAAAGHDVLFIGTD 453
OY 553 SGVLYKATIALQAGSAPREVEVLEELQVFPPTIMEISVKKRMILYVGSRLGVAOURL 612
DB 454 VGVLKATIASVPKGRSPSAGEGLLELHVEDSAVTSMSKSHQLYVASSNAQIAL 513
OY 613 HOCETGTACAECCCLARDPCAMDGASTHYRSLKRRFRFODIRHGNPALQCGSOG 672
DB 514 HRCAGHGRVCTECCCLARDPCAMDGASTHYRSLKRRFRFODIRHGNPALQCGSOG 672
OY 673 EEAVALYATMYGTEHNSSTLECPKSPQAAVRLRLDDEGDDVYKDERVLTHERG 732
DB 573 P-----ALHEHKEVGEVGSASFLECEPRSLQARVEMTFORACVTAHTOVLEERTERTARG 628
OY 733 LLERLSFFDAGTYCTTLEHGFSGOYVRLALVIVASOLDNLFP-PEPPPEPPARGGL 791
DB 629 LLERLSFFDAGTYCTTLEHGFSGOYVRLALVIVASOLDNLFP-PEPPPEPPARGGL 791
OY 792 ASTPKAMYNKIIOLI-----GRANLPVDEYCEYRWCRGTTSCSC-FRSRSGQAR 844
DB 685 -----PKLMTYDFDLQVLEPPGGGSANSLRM-----CRPQALQSLPSLRKGRNR 731

RESULT 14

AA27127
ID AAY27127 standard; Protein; 777 AA.

AC AAY27127;
XX 14-SEP-1999 (first entry)

DE Human brain tissue-derived polypeptide (clone OM007).

KM Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
KM recombinant; diagnosis; treatment.

OS Homo sapiens.

Key Location/Qualifiers
FT Peptide 1..36
FT Protein /note="signal peptide"
FT /note="mature protein"

PN W09933873-A1.

PD 08-JUL-1999.

PF 25-DEC-1998; 98WO-JP05952.

PR 26-DEC-1997; 97JP-0358811.

PA (ONOX) ONO PHARM CO LTD.

PI Fukushima D, Shibayama S, Tada H;

DR WPI; 1999-419088/35.

DR N-PSDB; AAX89112, AAX89113.

PT New adult human brain tissue-produced polypeptides useful for
PT diagnosis and treatment

PS Claim 1; Page 36-39; 86pp; Japanese.

CC The invention provides polypeptides (AAY27127-Y27133) produced by human
CC adult brain tissue, human bone marrow or a human umbilical cord venous
CC endothelial cell. Host cells transformed with vectors comprising the
CC nucleic acids encoding the polypeptides are used for the recombinant
CC expression of the polypeptides. The polypeptides can be used in
CC diagnosis, treatment and basic studies, with wide applications in
CC treatment depending on the activity to be aimed at. Sequences
CC AAX89112-125 represent nucleic acids encoding the polypeptides.
XX

SO Sequence 777 AA;

Query Match
Best Local Similarity 36.3%; Score 1721; DB 20; Length 777;
Matches 346; Conservative 144; Mismatches 229; Indels 48; Gaps 17;

OY 122 SVRLRLSRLDLSNRSALIFGPGSGLNQLAMYLDEYRDLRPLFGGLDALYSLRLDAMP 181
DB 122 SVRLRLSRLDLSNRSALIFGPGSGLNQLAMYLDEYRDLRPLFGGLDALYSLRLDAMP 181
OY 41 NIPRLKLTGKDLNLSNCPFLGSGSEGLDFTLLDDEBERGLLIGAKDHLFLSLVDL 100
DB 41 NIPRLKLTGKDLNLSNCPFLGSGSEGLDFTLLDDEBERGLLIGAKDHLFLSLVDL 100
OY 182 DREVLMPPOGQRECVRKGRDPLTECANFVRLQPHNTHLACGAFOPCALITY 241
DB 182 DREVLMPPOGQRECVRKGRDPLTECANFVRLQPHNTHLACGAFOPCALITY 241
OY 101 NKKLTPAAKEREVELCKLAGKDANTECANFVRLQPHNTHLACGAFOPCALITY 160
DB 101 NKKLTPAAKEREVELCKLAGKDANTECANFVRLQPHNTHLACGAFOPCALITY 160
OY 242 G-HNGEHLHLEPSSVSGRGCHPSPASTFIDELIYGLDDELREAMIFRSGG 300
DB 242 G-HNGEHLHLEPSSVSGRGCHPSPASTFIDELIYGLDDELREAMIFRSGG 300
OY 161 GYKEDLIFKLDTRNLSGLKCPFPQGFASVMTDELYSGTASDLGKDTATFSLG 220
DB 161 GYKEDLIFKLDTRNLSGLKCPFPQGFASVMTDELYSGTASDLGKDTATFSLG 220
OY 301 P-----RPAIRD-SDQSLHDPFVMAARIPENDODNDKVFYEFSETPSPDGSNHTV 356
DB 301 P-----RPAIRD-SDQSLHDPFVMAARIPENDODNDKVFYEFSETPSPDGSNHTV 356
OY 221 PTHDHNHITRDISHYNLWNGAKFICFTFIPDTIYNPDCKIYFFERES--SOEGSTDKTI 278
DB 221 PTHDHNHITRDISHYNLWNGAKFICFTFIPDTIYNPDCKIYFFERES--SOEGSTDKTI 278

[illegible]

Thu Oct 10 09:35:04 2002

us-09-813-290-2.1.rge

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
October 9, 2002, 19:18:09 ; Search time 2790 Seconds
(without alignments)
6562.982 Million cell updates/sec

Title: US-09-813-290-2
Sequence: 1 MACALAGKVPKMSRPVNHK.....KKMSRVHAEHNTPREVEAT 875
Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 10.0, Ygapext 0.5
Xgapop 6.0, Ygapext 7.0
Xgapop 6.0, Delcost 7.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DEV=tbl
-Q/cgn2.1/USPTO.spool/US09813290/runat.09102002.094645.19122/app_query.fasta.1.1031
-Q/cgn2.1/USPTO.spool/US09813290/runat.09102002.094645.19122/app_query.fasta.1.1031
-DB=GenEmbl -QFMT=fastp -SUFFIX=rig -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0.5
-UNITS=DITS -START=1 -END=1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=13 -MODE=LOCAL
-OUTFMT=pcol3290.ecgn.1.1.1828-efunat.09102002.094645.19122 -DCPU=6 -ICPU=3
-USER=US09813290.ecgn.1.1.1828-efunat.09102002.094645.19122 -WAIT -LONGLOG -DEV.TIMEOUT=120
-NO_XLPXY -NO_LMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*

28: em_vl:*
30: em_htg_inv:*
31: em_htg_inv:*
32: em_htg_inv:*
33: em_htg_inv:*
pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	4746	100.0	2628	6	AX253545	AX253545 Sequence
2	4746	100.0	3568	6	AX253549	AX253549 Sequence
3	4215	88.8	2349	6	AX253547	AX253547 Sequence
4	4215	88.8	4700	6	AB029496	AB029496 Homo sapi
5	4064.5	85.3	2340	6	AX155170	AX155170 Sequence
6	3913	82.4	4469	2	AK024425	AK024425 Homo sapi
7	3122.5	65.8	214046	2	AC094021	AC094021 Homo sapi
8	3110.5	65.5	123943	9	AC006208	AC006208 Homo sapi
9	3102.5	65.4	123943	6	AX155172	AX155172 Sequence
10	2282.5	48.1	229706	2	AC095672	AC095672 Rattus no
11	2056	43.3	2898	10	AF034744	AF034744 Mus muscu
12	2043.5	42.8	2615	5	AF022947	AF022947 Gallus ga
13	2033.5	42.5	6474	10	AB002329	AB002329 Human mRN
14	1955	41.2	3988	10	MM293948	MM293948 M. musculus
15	1955	41.2	4466	10	MM293947	MM293947 M. musculus
16	1904	40.1	3263	5	GG002528	GG002528 Gallus gall
17	1846.5	38.9	5952	10	MS002528	MS002528 Mus musculu
18	1846	38.8	2319	10	RN511001	RN511001 Mus musculu
19	1843	38.8	2319	10	RN511001	RN511001 Mus musculu
20	1840	38.8	2319	10	RN511001	RN511001 Mus musculu
21	1835.5	38.7	2913	10	MMRNASEMD	MMRNASEMD Sequence
22	1835.5	38.6	2709	6	AX207154	AX207154 Homo sapien
23	1834	38.6	2230	11	G31703	G31703 SMS1973 Et
24	1830	38.6	2230	11	G31703	G31703 SMS1973 Et
25	1830	38.6	2230	11	G31703	G31703 SMS1973 Et
26	1830	38.6	2230	11	G31703	G31703 SMS1973 Et
27	1830	38.6	2230	11	G31703	G31703 SMS1973 Et
28	1797	37.9	2715	5	GS028240	GS028240 Gallus gall
29	1779	37.5	2919	5	AF086761	AF086761 Homo sapien
30	1773	37.4	2919	5	AF086761	AF086761 Homo sapien
31	1769	37.3	2825	9	BC013975	BC013975 Homo sapi
32	1755	36.5	2825	9	AF217991	AF217991 Homo sapi
33	1733.5	36.3	2825	9	AF217991	AF217991 Homo sapi
34	1721	36.3	3880	6	E27342	E27342 Novel poly
35	1698.5	35.8	1998	10	MUSSEMAITI	MUSSEMAITI Sequence
36	1673	35.3	2499	10	BC010976	BC010976 Mus muscu
37	1667.5	35.1	3552	9	HSU38276	HSU38276 Human semap
38	1667.5	35.1	2872	10	MMRNASEMA	MMRNASEMA Sequence
39	1664.5	35.0	2265	10	HS033920	HS033920 Mus muscu
40	1659.5	34.7	2719	10	AF080091	AF080091 Mus muscu
41	1644.5	34.1	2765	10	AF080090	AF080090 Mus muscu
42	1618.5	33.9	2477	10	BC009113	BC009113 Homo sapi
43	1568	32.9	2336	5	AF022946	AF022946 Mus musculu
44	1561.5	32.9	2336	5	AF022946	AF022946 Mus musculu
45	1561.5	32.9	2336	5	AF022946	AF022946 Mus musculu

ALIGNMENTS

RESULT 1
AX253545
LOCUS AX253545
DEFINITION Sequence 1 from Patent WO0170806.
ACCESSION AX253545
VERSION AX253545.1 GI:16074048
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

source	1. 3568	"/db_xref="taxon:9606"	680 t
BASE COUNT	633 a	1099 c	1156 g
ORIGIN			
Alignment Scores:	7,96e-182	Length:	3568
Pred. NO.:	4746.00	Matches:	875
Score Similarity:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	6	Gaps:	0
DB:			
US-09-813-290-2 (1-875) x AK253549 (1-3568)			
QY	1 MetAlaGyAlaLeuAlaGlyLyValPheProMetGlySerTrpProValTrpHisLys	614	
Db	555 ATGGCTGTCCTCCCTACTCTCGAAGAGTGTCTCCATGGGGAGCGTGGCCAGTGTGCACAA	674	
QY	21 SerLeuHisTrpAlaAsnLysValGluGlyLysAlaGlyValGlyArgGlnGlyProSer	674	
Db	615 AACCTGCACCTGGCCACACAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	734	
QY	41 LeuLeuLeuSerSerAlaProLeuProAlaGlnAspTrpValGlnProLeuProTyrLys	734	
Db	675 CTCCTCTCTCTCTCCGCCCTCTCCGCCACAGAGTGGTGGAGCCACTGCTTAAG	794	
QY	61 TrpTrpProGlnGlySerArgAlaAsnTyrAsnArgArgProAlaGlyProGlnGlyLys	794	
Db	735 TGGTGGCTGGTGGACACACACAAACCAACACACACACACACACACACACACACACAC	854	
QY	81 SerAlaGlyArgArgGlnArgGlySerProGlnPheProSerMetAlaProSerAlaTrp	854	
Db	795 TGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	914	
QY	101 TACGATPLeuLeuGlyLysLeuLeuLeuHisGlyLysSerGlyProSerProGly	914	
Db	121 ProSerValProArgLeuArgLeuSerTyrArgAspLeuSerAlaAsnArgSerAla	974	
QY	915 CCCAGTGGCCCCCGCTGGCGCTCTCTCTCCAGACCTCTGCTCCCAACCGCTTGGC	1034	
Db	141 TlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGlnTyrArg	1034	
QY	975 ATCTTCTGGGCCCCCAGGGCTCTCTCAACCTCCAGGCCATGTACCTAGATGTACCA	1094	
Db	161 AspArgLeuPheLeuGlyLysLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaTrp	1094	
QY	1035 GACCGCTCTTCTGGGTGGCTGGACCGCTTACTCTTGGGCTGGACAGGACTGG	1154	
Db	181 ProAspProArgGluValLeuTrpProProGlnProGlnArgGlnGluLysValArg	1214	
QY	1095 CCGAGATCCCGGGAGGCTCTGTGGACCGGACGACGACGACGACGACGACGACGACG	1214	
Db	201 LysGlyArgAspProLeuTrpGluCysAlaAsnPheValArgValLeuGlnProHisAsn	1274	
QY	1155 AAGGAGAGATCTCTTGGACAGAGGACCCAACTTGTGGGTGTACACACTCCACAA	1274	
Db	221 ArgTrpHisLeuLeuLacGlyLysTrpGlyValPheGlnProTrpHisCysAlaLeuIle	1334	
QY	1215 CGAGACCCACCTGCTAGCTGTGGCATGGGGCTTCCAGGCCACCTGTGGCTATCACA	1334	
Db	241 ValGlnHisArgGlyGlnLysHisValLeuHisLeuGlnProGlySerValGlnSerGlyArg	1394	
QY	1275 GTGGGCGACCGTGGGAGACATGTCTCCACTCGAGCCGAGCGGAGGTGGAGAAATGGCCGG	1454	
Db	261 GlyArgCysProHisGlnProSerArgProPheAlaSerThrPheIleAspGlyLysLeu	1514	
QY	1335 GGGGCGGCGCTCGAGAGCCGAGCGCTCTTGGCAGCACCTCTCATAGCGGGAGCTG	1574	
Db	281 TyrThrGlyLeuThrAlaAspPheLeuGlyArgGlnAlaMetIlePheArgSerTyrLys	1634	

D	b	2475	ACCCACTAGCGGCCCTTGGCAAGCGCCGCGTTCCGCGCGAGGACAATCCGGCAGGC	2534
OY	661	AsnProAlaLeuGlnCysLeuGlnVgIsSerGlnIuGluAlaValGlyLeuValAla		
D	b	2535	AACCCCTGCCCTTCAGTGCTGTGGGCAGAACGAGGAAGAAGAGCAGTGGACTGTGGCA	680
OY	681	AlaTrpMetValTyrglyThrGlnHisSsnSerThrPheLeuGlnCysLeuProLysSer		
D	b	2595	GCCACCATGGCTCACGGCAAGGACACAAATACACTTCTCGAGAGTCCCGCCAACTCT	2594
OY	701	ProgInAlaAlaValArgPheLeuGlnAargProGlyLysPrGlnLyProaspGlnVal		
D	b	2655	CCCCAGCTGGCTGTCCGCTGGCTTCTTCGACAGGCGAGGGATGAGGGGCTGTACCAGTG	2654
OY	721	LysThrAspGlnArgValLeuHisThrGlnIarGlyLeuLeuPheArgValArgLeuSerArg		
D	b	2715	AAGAGCAGCAGCGACGACTTGCACACGAGGCGGGGGCTGTCTTCCAGAGCTTAAGCCT	2714
OY	741	PheAspAlaGlyThrTyrrThyrCysThrThrLeuGlnHisGlyPheSerGlnTrpVal		
D	b	2775	TTCGATCCGGGCACTTACACTGTGACACCACTCTGGAGAGCTTCTTCCAGAGCTTAAGCCT	2774
OY	761	ArgLeuAlaLeuValValIleValAlaSerGlnLeuAspAsnLeuPheProProLysPro		
D	b	2835	CCCCTGGCTGT	2834
OY	781	LysProGlnGluProProAlaArgGlyGlyLeuAlaSerThrProProLysAlaTrpTyr		
D	b	2895	AAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2894
OY	801	LysAspIleLeuGlnLeuIleGlyPheAlaAsnLeuProArgValAspGlyTyrCysGln		
D	b	2955	AAGGACATCTCTCACTCACTATTGGCTTGGCCCAACTGCGCCGGGTGATGATGATGATGAT	2954
OY	821	ArgValITrPCysArgGlyThrThrGlnCysSerGlyCysPheArgSerArgSerArgGly		
D	b	3015	CCCGT	3014
OY	841	LysGlnAlaArgGlyLysSerTrpAlaGlyLeuGlnLeuGlnGlyLysMetLysSerArg		
D	b	3075	AACGAGGCCAGGGGCAAGAGCTGGGACAGGCTGGAGCTAGCAAGAGATGAAGAGCCGG	3074
OY	861	ValHisAlaGlnHisAsnArgThrProArgGlnValGlnLysThr		
D	b	3135	GTGCATGCGGACCAATCGAGCCCGCGGAGGTGAGGCCACG	3134
RESULT 3				
LOCUS	AX253547			
DEFINITION	Sequence from Patent WO0170806.	2349 bp	DNA	linear PAT 10-OCT-2001
ACCESSION	AX253547			
VERSION	AX253547.1			
KEYWORDS	GI:16074049			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	(bases 1 to 2349)			
TITLE	Maize, D.W., Wilgenbusch, N.L., Turner, C.A., Hilborn, E., Wang, X.,			
JOURNAL	Human secreted proteins and polynucleotides encoding the same			
FEATURES	Lexicon Genetics Incorporated (US)			
Source	Location/Qualifiers			
	1..2349			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	ORIGIN	406 a	747 c	753 g
				442 t
				1 others

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 88.81% Indels: 0
 DB: 6 Gaps: 0

US-09-813-290-2 (1-875) x AX253547 (1-2349)

94 MetAlaProSerAlaTrpAlaIleCysTrpLeuLeuGlyLeuLeuLeuHisGlyGly 113
 1 ATGGCCCCCTGGCCCTGGGCAATTCCTGCTAGAGGGGCTCTCTCCATGGGGGT 60
 114 SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTyrArgSpleu 133
 61 AGCTCTGGCCCAACCCCGCCGAGTGTCCCGCTGGGCTCTCTAACAGACCTTC 120
 134 LeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnGlnAla 153
 121 CTGTCTGCACACCGCTCTGCCATCTTCTGGGCCCCCAGGGCTCTGAACTCCAGGCC 180
 154 MetTyrLeuAspGlyTyrArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSer 173
 181 ATGTACCTAGATGAGTACCGAGACCGCTCTTCTGGGTGGCTGAGCGCCCTTACTCT 240
 174 LeuArgLeuAspGlnAlaIleTrpProAspProArgGlyValLeuTyrProProGlnProGly 193
 241 CTCGGCTGGACAGGCGATGGCCAGATCCCGGGAGTCTGTGGCCACCGCAGCAGAGA 300
 194 GlnArgGlyGlyCysValArgGlyGlyArgAspProLeuThrGlyCysAlaAsnPheVal 213
 301 CAGAGGAGAGAGTGTGTGCAAGGAGAGAGATCTTTACAGAGTGGCCCAACTTCTGTG 360
 214 ArgValLeuGlnProHisAsnArgThrHisLeuLeuValAcysGlyThrGlyAlaPheGln 233
 361 CGGGCTGACAGCTCTCAACCCGACCCAGCTGCTAGCTGGGCACTGGGGCTTCCAG 420
 234 ProThrCysAlaLeuIleThrValGlyHisArgGlyGlyValLeuHisLeuGlnPro 253
 421 CCCACCTGGCCCTCATACAGATTGGCCACCGCTGGGAGCATGTGGCTCCACCTGAGCC 480
 254 GlySerValGlySerGlyArgGlyArgGlyProHisGlyProSerArgProPheAlaSer 273
 481 GCGAGGTGGAAAGTGGCGGGGGGGGCTCCACGAGCCAGCCGCTCTTGGCCAGC 540
 274 ThrPheIleAspGlyGlyLeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGlnAla 293
 541 ACCTTCATAGACGGGAGGCTGACAGGGTCTCAGCTGCTGACTTCTCTGGGGGAGAGGCC 600
 294 MetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAspSerAspGlnSerLeu 313
 601 ATGATCTTCGAAAGTGGAGGTCTCTCGGCCAGCTCTGCTCCGACTCTACCAAGATCTC 660
 314 LeuHisAspProArgPheValMetAlaAlaArgIleProGlnAsnSerAspGlnAspAsn 333
 661 TTGCGACGACCCCGGTGTGTGATGGCCCGCGGATCCCTGAGAACTCTGACCGAGCAAT 720
 334 AspGlyValTyrPhePhePheSerGlyThrValProSerProAspGlyGlySerAsnHis 353
 721 GACAGGGTCTACTTCTTCTCTCGAGAGCGCTCCCTCGCCGAGTGGTGGCTCGAACACT 780
 354 ValThrValSerArgValGlyArgValCysValAsnAspAlaGlyGlyGlnArgValLeu 373
 781 GTACACTGACCGCGCTGGCGCGCTGCGTGGATGATGGGGGGGCGACCGGTGCTG 840
 374 ValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGly 900
 841 GTGAACAAATGAGCACTTCTCCAGAGCCAGAGCTGTGCTCGGTGGCGCGCTCTG 960
 394 GlyAlaGlyThrHisPheAspGlnLeuGluAspValPheLeuLeuTrpProLysAlaGly 413
 901 GGTGCGGAGACCCCTTGTGACAGCTAGAGATGTGTCTCTGTCGCCCAAGCGCGGG 960
 414 LysSerLeuGlyValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAla 433

961 AAGACCTCGAGGTGTACGGGCTGTCTGACGACCGGTGACGTCGCTGTCCAGGGCTTCGCC 1020
 434 ValCysValTyrHisMetAlaAspIleTrpGlyValPheAsnGlyProPheAlaHisArg 453
 1021 GTGTGTGTACCAATGAGCAATCTGCGAGAGGTTCACAGGGGCTTGGCCACCGCA 1080
 454 AspGlyProGlnHisGlnTrpGlyProTyrGlyGlyValProPheProAspProGly 473
 1081 GATGGGCTTCAGACCAAGTGGGGGCTTATGGGGGCAAGTGGCTTCTCTGCGCTTGGC 1140
 474 ValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyr 493
 1141 GTGTGCCCCAGCAAGATGACCCGACAGCCACAGAGGAGCTTTGGCAGCACAAGGACTAC 1200
 494 ProAspGlyValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgPro 513
 1201 CCAGATGAGGTGTGCAAGTTTGGCCGAGCCACCCCTCATGTCTGTGGCTGTGGGGCT 1260
 514 ArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnLeuHisGlnIleVal 533
 1261 CGACATGGCCGCGCTGTCTGTGTCAGAGCCACCTGGCCACAGCTACACGATCTGT 1320
 534 ValAspArgValGlyAlaGlyAspGlyThrTyrAspValIlePheLeuGlyThrAspSer 553
 1321 GTGGACCGGTGGAGGAGGAGATGAGCACTACAGATGCTATTTCTGGGGAGTGACTCA 1380
 554 GlySerValLeuLysValIleAlaLeuGlnAlaGlyGlySerAlaGlyProGlyGlyVal 573
 1381 GGGTCTGTGCTCAAAATGATCTGCTCCAGGAGGGGCTGAGCTGAACTGAGAGAGTG 1440
 574 ValLeuGlyGlyLeuGlnValPheLysValProThrProIleThrGlyMetGlyLysSer 593
 1441 GTCTGGAGGAGCTCCCGGTGTAAAGTGTCCACACCTATACCAAGAAATGAGATCTCT 1500
 594 ValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHis 613
 1501 GTCAAAAGCAAAATGCTATACGTGGCTGTGGCTGTGGCTGTGGCCGCTGGCGTGCAC 1560
 614 GlnCysGlyThrTyrGlyThrAlaCysAlaGlyCysCysLeuAlaArgAspProTyrCys 633
 1561 CATGTGAGACTTACGGCAGCTGCTGTGACAGATGCTGCGCCGCGGAGCCCATCTGT 1620
 634 AlaTrpAspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArgAspArg 653
 1621 GCCTGGATGTGCTCTCTGTACCCACTACCGCCACCTTGGCAGAGCGGCTTCCGC 1680
 654 ArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlyGlnSerGlnGlyGlu 673
 1681 CGGAGAGACATCCGGCAGCGCAACCTGCTGCAAGTGGCTGGCCAGAGCCAGAGACA 1740
 674 GlnAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGlnHisAsnSerThrPhe 693
 1741 GAGCGATGGGACTTGTGGCAGCCACCATGTGTACGGCAGGAGCAAGAAATGACACTTC 1800
 694 LeuGlyCysLeuProLysSerProGlnAlaAlaValArgTrpLeuLeuGlnArgProGly 713
 1801 CTGAGATGCTCCGCCAATCTCCCAAGCTGTGCTGGCTGTGGCTGTGGCAGAGCCAGGG 1860
 714 AspGlyGlyProAspGlnValLysThrAspGlyArgValLeuHisThrGlnArgGlyLeu 733
 1861 GATGAGGGGCTGACAGAGTGAAGAGGAGGAGGAGGAGCTTGGCAGCAGCGAGCGGGCTG 1920
 734 LeuPheArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrThrLeuGlnHis 753
 1921 CTGTTCGCAAGGCTTACCGCTTTCGATGGGGGACCTACCTGACCACTCTGGAGCAT 1980
 754 GlyPheSerGlnThrValValArgLeuAlaLeuValIleValAlaSerGlnLeuAsp 773
 1981 GCGTCTCCAGACTGTGGTCCGCTGCTGTGTGTATGTATGGCTGCACAGCTGAGAC 2040
 774 AsnLeuPheProProGlyProLysProGlyGlyProProAlaArgGlyGlyLeuLysSer 793
 2041 AACCTGTCTCTCGGAGCCAAAGCCAGAGAGCCCGGAGAGCGCTGGCTTCC 2100

```

QY 794 ThrProLysAlaIlePtyrLysAspIleLeuGlnIleGlyPheAlaIlePro 813
Db 2101 ACCCCAGCCAGCCGCTGGTACAGACATCTGCAGCATCTGCTTGGCCAACTGGCC 2160
QY 814 ArgValAspGluTyrCysGluArgValTrrPcysArgGlyThrThrGluCysSerGlyCys 833
Db 2161 CGGGGAGTGAAGTACTGTGAGCGCGGTGTGTGTGCAAGGGGACACAGGAATGCTCAGGCTGC 2220
QY 834 PheArgSerArgSerArgGlyGlnAlaArgGlyLysSerTrrPalaGlyLeuGlnLeu 853
Db 2221 TTTCCGAGACCCGAGCCGCGGAGACAGCCAGGCGCAAGAGCTGGCAGGCGCTGAGCTTA 2280
QY 854 GlyLysLysMetLysSerArgValHisAlaGlnHisAsnThrThrProArgGluValGlu 873
Db 2281 GGCAGAGAGATGAAGACCCGCGGTGCATGCCGACACATGAGAGCCCGCGGAGTGTGAG 2340
QY 874 AlaThr 875
Db 2341 GCCACG 2346

RESULT 4
AB029496 4700 bp mRNA linear PRI 07-JUL-2000
LOCUS DEFINITION AB029496 Homo sapiens mRNA for semaphorin sem2, complete cds.
ACCESSION AB029496.1 GI:8978201
VERSION semaphorin sem2.
KEYWORDS semaphorin sem2.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,
Miyajima, N. and Saito, T.
Human semaphorin
2 (bases 1 to 4700)
Published Only in Database (2000) In press
Seki, N., Hattori, A., Hayashi, A., Kozuma, S., Muramatsu, M.,
Miyajima, N. and Saito, T.
Direct Submission
Submitted (01-JUL-1999) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group, Inage-Ku Anagawa
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail: t_saito@nirs.go.jp,
Tel: 81-43-201-3135, Fax: 81-43-251-9818)
FEATURES
Source Location/Qualifiers
1..4700
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2349
/feature="sem2"
1..2349
/feature="sem2"
/codon_start=1
/product="semaphorin sem2"
/protein_id="BAA98132.1"
/db_xref="GI:8978201"
/translation="MAPSAMATCWLIGGULLHGGSGPSPVRLRLSYRLISAN
BSAIFGPGSLNQAMLYDERDRLTGGIDALYSLLDAMPDPREVLPPOGR
ECSVKGKGDPLTECANPRAVLPDPNRLHLAGCGAPQPTALLTVHGSHVLEP
GSVEGGRGRCHPERSRPAFTFIDELTYGLTADFGREANIFELYSRPRRLHLEP
SLDHPREVMAARLPENSDDNDKVFEEFSETPSPGSGNHVTVSVAGVANDAG
OVFQPCVYTHMADINVEFNGPFHNRDGPQHGKVPPEPQVPSKMTAQR
PSTGKDPDEVLQFAHNPDMFVPRPRGRVLTHTLAQDLQIVDRAEAGT
IDVIFLTDGSLVKVAILQAGSAPREVEYLLELDVFKVPTITMEISVKROMLY
GSRGLQALRHOCCEYTGACACCCCLARDYCAMDASCHYHPSLGRFRRLDRIH
GNALQCLGSOEEAVGLVAATVYGTENSTFLCPLKSPQAAVRLWLTQVSDGCP
DQVTEBERVLPHTERGLRLRSLRFDAGTYTCTLEHGFQSTVRAVRLVYASQDNL
FPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEP
FRSRSGKARSGMAGLEGRKMSRVAAEHNRPREVEAT"
BASE COUNT
972 a 1307 c 1467 g 954 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1 79e-160
Score: 4215.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 88.81%
DB: 9
Gaps: 0

US-09-813-290-2 (1-875) x AB029496 (1-4700)
QY 94 MetAlaProSerAlaTrrPalaIleCysTrrLeuLeuGlyGlyLeuLeuHisGly 113
Db 1 ATGGCCCCCTGGCTGGCCATTTGTGTGGCTGTGAGGGGCTCTCTCTCATGGGGT 60
QY 114 SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTyrArgAspLeu 133
Db 61 AACTGTGCCCCAGCCCGCGCCAGTGTGCCCGCTTCCGCTCTACCGAGACTC 120
QY 134 LeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAla 153
Db 121 CTGTCTGCCAACCGCTCTGTCATCTTCTGGCCCCCAGGGCTCCCTGAACTCCAGGCC 180
QY 154 MetTyrLeuAspGlyTrrArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSer 173
Db 181 ATGTACTTATGATGATGACGAGACCCGCTCTTCTGGGTGGCTGAGACCGCTTACTCT 240
QY 174 LeuArgLeuAspGlnAlaTrrProAspProArgGluValLeuTrrProGlnProGly 193
Db 241 CTGGGCTGACAGCAGCATGCGCAGATCCCGGAGGTCTGTGCCACCGCAGCAGCA 300
QY 194 GlnArgGlnGlyCysValArgGlyGlyArgAspProLeuThrGluCysAlaAsnPheVal 213
Db 301 CAGAGGAGAGACTGTCTTCTGAAAGGAGAGATCTTGTACAGAGTGGCCCACTTCTGTG 360
QY 214 ArgValLeuGlnProHisAsnArgThrHisLeuLeuLacGlyGlyThrGlyAlaPheGln 233
Db 361 CGGGGTCTACAGCTCTCACACACCGGACCCACTGTACTGTGGGCTTCCAG 420
QY 234 ProThrCysAlaLeuIleThrValGlnHisArgGlyGlyHisValLeuHisLeuGlyPro 253
Db 421 CCCACCTGTGCTCATCAGATGTTGGCCACCGTGGGAGCATGTGCTCCACCTGAGGCT 480
QY 254 GlySerValGlnSerGlyArgGlyArgCysProHisGlyProSerArgProPheAsnSer 273
Db 481 GGCAGTGTGAAAGTGGCGCGGGCGGTGCTCCACGAGCCACCGCTCTTGTCCAGC 540
QY 274 ThrPheAlaSerGlyGlyLeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGlnAla 293
Db 541 ACCTTATAGAGGGAGGAGCTGTACAGGGGTCTCACTGCTGACTTCTGGGGCAGAGGCC 600
QY 294 MetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAspSerAspGlnSerLeu 313
Db 601 ATGATCTTCCGAGTGGAGGTGCTCTGGCCACTCTGCTTCCGACTCTGACACAGACTCTC 660
QY 314 LeuHisAspProCysPheValMetAlaAlaArgIleProGluAsnSerAspGlnAspAsn 333
Db 661 TTGCAGCAGCCCGCTGTTGTGATGGCCCGGAGTCCCTGAGAACTCTGACAGACAAT 720
QY 334 AspLysValTyrPhePhePheSerGlyThrValProSerProAspGlyGlySerAsnHis 353
Db 721 GACAGAGTGAATCTTCTCTCTGAGAGCGGTCCCTCGCCAGTGTGGTGTGAAACAT 780
QY 354 ValThrValSerArgValGlyArgValCysValAspAlaGlyGlyGlnArgValLeu 373
Db 781 GTCACTGTACGCGCTGGCGCGCTGCTGTAATGATGCTGGGGGCCAGGGGTGTCTG 840
QY 374 ValAsnLysTrrSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGly 840
Db 841 GTGAACAAATGAGACACTTCTCTCAAGCGCAGGCTGTGCTGCTGCTGCTGCTGCTGCT 900
QY 394 GlyAlaGluThrHisPheAspGlnLeuGlnAspValPheLeuLeuTrrProLysAlaGly 413

```

Db 501 GGTCGCGAGACCACTTTGACCAAGTATGTTCTCTGTCGCCCCAGGCCGG 960
 Oy 414 LysSerLeuGluValThrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAla 433
 Oy 961 AAGACCTCGAGAGTATGAGCGCTGTCAGCAACCGTACAGTGGCTGTTCAGAGGCTTCGCC 1020
 Oy 434 ValCysValIleThrHisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisArg 453
 Db 1021 GTCGTGTGTACCATGATGGACATCTGTGGAGGTTTTCACGGGGCCCTTTGCCACCGCA 1080
 Oy 454 AspGlyProGlnHisGlnTrpGlyProGlyGlyValProPheProAspProGly 473
 Db 1081 GATGGGCTCACACAGTGGGGGCTATGGGGGCAAGGTGCTCTCCCTCCCTCCGCGC 1140
 Oy 474 ValCysProSerLeuMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyr 493
 Db 1141 GGTGCCCCAGCAAGTATGACCCAGCAGGAGCGCTTTGGCAGCACAAGAGTATC 1200
 Oy 494 ProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgPro 513
 Db 1201 CCAGATGAGAGTGTGATGTTGCCAGGCCACCCCTCATGTTCTGTGGCTGTGGCGCT 1260
 Oy 514 ArgHisGlyArgProValLeuValIleThrHisLeuAlaGlnGlnLeuHisGlnIleVal 533
 Db 1261 CGCATGCGCCGCTCTCTCTGTCAGAACCCACCTGGCCAGAGCTACACAGATGCTG 1320
 Oy 534 ValAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThrAspSer 553
 Db 1321 GTCGACCGGCTGAGGAGCAGAGTGGAGCTACGATGCTATTTTCGAGGAGTACATCA 1380
 Oy 554 GlySerValLeuLysValIleAlaLeuGlnAlaGlySerAlaGluProGluGluVal 573
 Db 1381 GGGTGTGTGCTCAAGTATCATGCTCTCCAGCAGGGGGCTCAGCTGAACCTGAGGAATG 1440
 Oy 574 ValLeuGluGluLeuGlnValPheLysValProThrProIleThrGluMetGluLeuSer 593
 Db 1441 GTCTGTGAGAGAGCTCCAGGTGTTAAGTGCACACACTATCACGGAATGAGATCTCT 1500
 Oy 594 ValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHis 613
 Db 1501 GTCMAAAGGCAAAATGCTATAGTGGCTCTGCGGTGGGTGGCCAGCTGGCGCTGCAC 1560
 Oy 614 GlnGluGluThrTyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 633
 Db 1561 CAATGTAGACTTACGGCAGCAGCTGTGCAAGTGTGCTGCGCGGAGCCATACCTGT 1620
 Oy 634 AlaTrpAspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArgArgPheArg 653
 Db 1621 GCTGGGATGCTGCTCTCTATCCACTACCGCCCGCCCTTGGCAAGCGCGGTTCGCG 1680
 Oy 654 ArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlnGlnSerGlnGlu 673
 Db 1681 CGGAGGAGCATCCGGCAGCGGCAACCTGCTGCGAGTGTGCTGGCCAGAGCCAGGAAGA 1740
 Oy 674 GluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGluHisAsnSerThrPhe 693
 Db 1741 GAGCAGAGGGAGCTTGTGAGCAGCCACATGTGTACGCGCAGCAGGACACATATGACCTTC 1800
 Oy 694 LeuGluCysLeuProLysSerProGlnAlaValAlaArgTrpLeuLeuGlnArgProGly 713
 Db 1801 CTGGAGTCTCTGCCCAAGTCTCTCCAGGCTGTGTGGCTGTGGCTGTGGCAGAGGCCAGG 1860
 Oy 714 AspGluGlyProAspGluValLysThrAspGluArgValLeuHisThrGluArgGlyLeu 733
 Db 1861 GATGAGGGGCTGACAGAGTGAAGAGCGAGCGAGAGTCTTGACACAGGAGCGGGGCTG 1920
 Oy 734 LeuPheArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrTrpLeuGlnHis 763
 Db 1921 CTGTTCGCGAGGCTTACCGGTTTCATGTGGGCGACCTACACCTGCACTGTGGAGCAT 1980
 Oy 754 GlyPheSerGlnThrValValArgLeuAlaLeuValValIleValAlaSerGlnLeuAsp 773
 Db 1981 GGCTTCTCCACAGTGTGTGCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040

Oy 774 AsnLeuPheProGluProLysProGluGluProAlaArgGlyGlyLeuAlaSer 793
 Db 2041 AACCTGTCTCTCCGAGACCAACCGACAGAGAGCCCGGAGGAGGCTGTGCTTCC 2100
 Oy 794 ThrProLysAlaTrpTyrLysAspIleLeuGlnLeuIleGlyPheAlaAsnLeuPro 813
 Db 2101 ACCCCACCAAGGCTGTGTACAGACATCTGTGCTCATGTGCTGTGCTGTGCTGTGCT 2160
 Oy 814 ArgValAspGluTyrCysGluAlaArgValTrpCysArgGlyThrThrGluCysSerGlyCys 833
 Db 2161 CGGGTGTAGTACTGTGTGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
 Oy 834 PheArgSerArgSerArgGlyLysGlnAlaArgGlyLysSerTrpAlaGlyLeuGlnLeu 853
 Db 2221 TTCCGAGACCGGAGCGCGGGGAGCAGAGCCAGGCAAGAGCTGGGCGAGGCTGAGCTA 2280
 Oy 854 GlyLysLysMetLysSerArgValHisAlaGlnHisAsnArgThrProArgGluValGlu 873
 Db 2281 GGCAGAAGATGMAAGCGCGGTGATGCCGAGCACAATCGAGCCCGGAGGTGGAG 2340
 Oy 874 AlaThr 875
 Db 2341 GCCACG 2346
 RESULT 5
 AX155170 2340 bp DNA linear PAT 22-JUN-2001
 LOCUS AX155170
 DEFINITION Sequence 1 from Patent WO0140278.
 ACCESSION AX155170
 VERSION AX155170.1 GI:14536655
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2340)
 TITLE Human semaphorin zsmf-16
 AUTHORS Holloway, J.L. and Foley, K.P.
 JOURNAL Human semaphorin zsmf-16
 JOURNAL Patent: WO 0140278-A 1 07-JUN-2001;
 ZymoGenetics, Inc. (US)
 FEATURES
 source
 1..2340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..2340
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC42673.1"
 /db_xref="GI:14536656"
 CDS
 /translation="MAPSAAATGMLGILLHGGSSGSPSPSVRLRLSYGAMVRK
 PSTMMETTSRLLSANRSATPLPGDSLNLQAMLYDEYRDLFLGGDLALSLRD
 QAMPDPRPTEPCANFVRVLPQHNRTHLACGTGAOPICALITVGHREHYHLRPSY
 ESGRGCPHEPSPRPASTFDIGELVLTLDLFLREMTFRSGSPRALRSDOSLL
 HDPRVMAARI PENSDDNDKVFYEFSETPSPDPGDSNHVSVRGVCVNDAGGRV
 LVNKSFTFLKARLYCSVPGAGETHTDQLEDFLLPMPKSKSLVLAFTSVAVFG
 GFAYCVYHMDIIVENVGPRFPAHRDGPQHWGPGYGGKFPFGYCPSPKMTAQPGRPG
 STDYQDEYVQPARHRLMMPVRPRRGRPVLYLTHLAQOLHQLVVDVEAEDGTYDY
 FLVTDGSSVLYKVIADAGSSAPREYVRLVLEIYQFKVPTTEMEISVKRMLYVGR
 LGVAQLRLHQCETVGTACACCLARDYCAMDASCTHYRSPQAVRMLDRPDEGDOY
 ALQCLGSOEAEAGVLAATVAGTENSFTLECLPSPQAVRMLDRPDEGDOY
 KTDREYVLTREGLFRSLRFDAGTCTTLEHFPSPQAVRMLDRPDEGDOY
 EPKPEEPARGGLASTPDKAWKDILDIGANLPVDEYCEYRWCNRTGTCSCGFS
 RSKCKORGSMSGLEIGKKMSVHVEHNTPREVENT"
 BASE COUNT 406 a 748 c 743 g 443 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,18e-154 Length: 2340
 Score: 4046.50 Matches: 760
 Percent Similarity: 95.00% Conservative: 0
 Best Local Similarity: 95.00% Mismatches: 1
 Query Match: 85.26% Indels: 39

```

DB: 6 Gaps: 2
US-09-813-290-2 (1-875) x AX155170 (1-2340)
OY 94 MetAlaProSerAlaTrrAlaIleCySTrrPLeuLeuGlyLeuLeuLeuHisGlyGly 113
DB 1 ATGGCCCCCTCGGGGCGGCGCATTTCTGCTCTAGGGGGGCGCTCTGCTCATGGGGGT 60
OY 114 SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTyrArgAsp 132
DB 61 AGCTCTGGCCCCGAGCGCGCCCAATGTGCCCCCGCTCGGGCTCTCTACCGAGAGCC 120
OY 133 -----LeuLeuSer 135
DB 121 ATGTCCGAAGCCCTTCAGACCATGTGATGAAACATTTTCCAGATACCTCTGTCT 180
OY 136 AlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyr 155
DB 181 GCCAAGCCCTCTGCTATCTTTTGGGCCCCCAGGGGCTCCCTGAACCTCCAGGCCATGTAC 240
OY 156 LeuAspGluTyrArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArg 175
DB 241 CTAGATGATGACCGAGAGCCGCTCTTTGGGGGCTGGAGCCCTCTACTCTCTGCGG 300
OY 176 LeuAspGlnAlaTrrProAspProArgGluValLeuTrrProGlnProGlyGlnArg 195
DB 301 CTGGACCAAGGATGGCCAGATCCCGGGAG----- 330
OY 196 GluGluCysValArgGlyGlyArgAspProLeuThrGluCysAlaAsnPheValArgVal 215
DB 331 -----ACAGAGTGGCCCACTTCGTGGGGGTG 357
OY 216 LeuGlnProHisAsnArgTrrHisLeuLeuAlaCysGlyTrrGlyAlaPheGlnProThr 235
DB 358 CTACAGCCTCAACAACGGAGCCACCTGCTAGCCTGGCACGCGGCGCTCCAGCCACAC 417
OY 236 CysAlaLeuLeuThrValGlyHisArgGlyGlnHisValLeuHisLeuGlnProGlySer 255
DB 418 TGTGCTCTCATCAGATGTGGCCAGCGTGGAGCATGTCTCTCAGCTGAGACCTGGCAGT 477
OY 256 ValGluSerGlyArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPhe 275
DB 478 GTGGAAAGTGGCGGGGGGCGGCTCCCTCAGACGCCAGCCGCTCTTCCAGCACCTTC 537
OY 276 IleAspGlyLeuLeuTrrGlyLeuThrAlaAspPheLeuGlyArgGlnAlaMetTie 295
DB 538 ATAGACGGGAGGTGTACAGCGGTCTCAGCTCTGCTCTGCGGCGAGAGGCCATGATC 597
OY 296 PheArgSerGlyGlyProArgProAlaLeuArgSerAspSerAspGlnSerLeuLeuHis 315
DB 598 TTCGGAAGTGGAGTCTCGGCGCAGCTCTGCTTCGACCTGACACAGAGTCTCTGCAC 657
OY 316 AspProArgPheValMetAlaAlaArgIleProGluAsnSerAspGlnAspAsnAspLys 335
DB 658 GACCCCGGTTTGTGATGGCGCGCCGATCCCTGAGACTCTGACACAGACATGACAAG 717
OY 336 ValTyrPhePhePheSerGluTrrValProSerProAspGlyGlySerAsnHisValThr 355
DB 718 GTCTACTTCTTCTTCGAGACGCTCCCTCGCCGATGTGGTGCATCAACATGTCACT 777
OY 356 ValSerArgValGlyArgValCysValAsnAspAlaGlyGlyGlnArgValLeuValAsn 775
DB 778 GTTCAGCCGGGTGGGCGCGCTGCTGTAATGATCTGCGGGCGCACGCGGTGTGTGAAC 837
OY 376 LysTrrSerThrPheLeuLysAlaArgLeuValLysSerValProGlyGlyAla 395
DB 838 AATGAGACACTTCTCTCAAGGCCAGCGCTGTGTGCTGCGCGCCGCTGTGTGTGTC 897
OY 396 GluTrrHisPheAspGlnLeuGluAspValPheLeuLeuTrrProLysAlaGlyLysSer 415
DB 898 GAGACCCACTTTTGACAGCTAGAGGATGTCTCTGCTGGGCCCAAGCGCGGAAAGGC 957
OY 416 LeuGluValTrrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
DB 958 CTCGAGGTGTACGGCGCTTCAGACACCGTCAAGTCCGCTTCAGAGGCTTCGCCGCTCT 1017
OY 436 ValTrrHisMetAlaAspIleTrrGluValPheAsnGlyProPheAlaHisArgAspGly 455
DB 1018 GTTACCACTTGGCAGACATCTGGGAGGTTCACAGGGGCGCTTGGCCACAGAGATGGG 1077
OY 456 ProGlnHisLeuTrrGlyProTrrGlyGlyLysValProPheProArgProGlyValCys 475
DB 1078 CTCACAGACCACTGGGGGCGCTTATGGGGGCAAGGTGCCCTTCCCTGCGCTGCTGTC 1137
OY 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTrrProAsp 495
DB 1138 CCCAGCAAGATGACCCGACAGCCAGAGCGGCTTTGGACACACCAAGACTACCCAGAT 1197
OY 496 GluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrrProValArgProArgHis 515
DB 1198 GAGGTCTGCTGAGTTTGGCCGAGCCACCCCTCATGTTCCTGGCTGCGCTGACAT 1257
OY 516 GlyArgProValLeuValLysTrrHisLeuAlaGlnGlnLeuHisGlnIleValValAsp 535
DB 1258 GGCCGCGCTGCTCTTTCAGAGCCACCTGGCCAGCAGCTACACAGATGCTGTGGAC 1317
OY 536 ArgValGluAlaGluAspGlyTrrTyrAspValIlePheLeuGlyTrrAspSerGlySer 555
DB 1318 CGCGTGGAGGAGAGGATGGAGCTAGATGTCTTTCTCTGGGGAGCTGACTGAGGTCT 1377
OY 556 ValLeuLysValIleAlaLeuGlnAlaGlyGlySerAlaGluProGluGluValLeu 575
DB 1378 GTGCTCAAGATATCTCTCTCCAGGAGGGGGCTCAGCTGAACTGAGGAGTGTCTTG 1437
OY 576 GluGluLeuGlnValPheLysValProThrProIleThrGluMetGluIleSerValLys 595
DB 1438 GAGGAGCTTCAGGTGTTTAAGTGCACACCTATACCCCAAAATGAGATCTGTCTGAAA 1497
OY 596 ArgGlnMetLeuTrrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnCys 615
DB 1498 AGGCAAAATGCTTAGCTGGCTCTCGGCTGGGGTGGCCAGCTGGCGGCTGCACCAATGT 1557
OY 616 GluTrrTyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAspProTrrCysAlaTrr 635
DB 1558 GAGACTTAGGGACAGTCCCTGTGCAGAGTGTCTGGCCCGGGAGCCATATCTGTGCTGG 1617
OY 636 AspGlyAlaSerCysThrHisTrrArgProSerLeuGlyLysArgArgPheArgGln 655
DB 1618 GATGTGCTCTCTGTACCCACTACCGCCCGACGCTTGGCAAGCGCGCTTCGCGGACAG 1677
OY 656 AspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlyGlnSerGlnGluGluAla 675
DB 1678 GACATCCGGCACGGCAACCTCCCTGCAAGTGTCTGGCCAGAGCCAGAAAGAGGCA 1737
OY 676 ValGlyLeuValAlaAlaThrMetValTrrGlyTrrGlnHisAsnSerThrPheLeuGlu 695
DB 1738 GTGGAGCTTGTGGAGCACCATGTGCTACGGCAGCAGACACAAATGACACTCTGTGAG 1797
OY 696 CysLeuProLysSerProGlnAlaAlaValArgTrrPheLeuGlnArgProGlyAspGlu 715
DB 1798 TGCCTGCCCAAGTCTCCCGAGGCTCTGTGCTGTGCTGTGAGAGGCGAGGATGAG 1857
OY 716 GlyProAspGlnValLysTrrAspGluArgValLeuHisThrCysLysArgGlyLeuPhe 735
DB 1858 GGGCTTGACCAAGGTGAACAGCAGCAGCAGAGCTTGCACACGAGAGGGGGGCTGTCTTC 1917
OY 736 ArgArgLeuSerArgPheAspAlaGlyTrrTyrTrrCysThrThrLeuGlnHisGlyPhe 755
DB 1918 CGAGGCTTACCCCTTTCGATGCGGGCAGCTACACCTGACACACTGTGAGAGCTGCTTC 1977
OY 756 SerGlnThrValAlaArgLeuAlaLeuValIleValAlaSerGlnLeuAsnLeu 775
DB 1978 TTCAGAGCTGTGGTCCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2037
OY 776 PheProGluProLysProGluGluProProAlaArgGlyGlyLeuLysSerThrPro 795

```



```

431 GlypHealAvalCysValTyrHisMetAlaAspIleThrPglUvalPheAsnGlyProPhe 450
|||||
841 GGCTTCGGCGTCTGTGTGTCACCATGCGACATCTGGGAGGTTTTCACGGGCGCTTT 900
451 AlaHisArgAspGlyProGlnHisGlnTyrPglYpProTyrGlyGlyValProPhePro 470
|||||
901 GCCCACCAGATGGGCGCTCAGCACCAGTGGGGCGCTATGGGGGCAAGTGCGCTTCCCT 960
471 ArgProGlyValCysProSerLeuMetThrAlaGlnProGlyArgProPheGlySerThr 490
|||||
961 CCCCTGGCGGTGTGGCCCGACAGATGACCGCAGCAGAGCGGCGCTTTTGGCAGCAGC 1020
491 LysAspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheThrPro 510
1021 AAGGACTACCCAGATGAGGTGTGTCAGTTTGGCCGAGCCACCCCTCATGTTCGGGCT 1080
511 ValArgProArgHisGlyLysProValLeuValLysThrHisLeuAlaGlnGlnLeuHis 530
|||||
1081 GTGGGCGCTCGACATGCGCGCGCTGTCTTCATGACAGCACCAGCGGCGCGCTACAC 1140
531 GlnIleValAlaAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGly 550
|||||
1141 CAGATCGTGATGACCGCGCTGGAGCGAGAGATGGACCTACATGATTCATTTCTGGGG 1200
551 ThrAspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlySerAlaGluPro 570
1201 ACTGACTCAGGCTGTGTCTCAAGTCATCGCTCTCCAGCAGAGGGGCTCAGCGAAGCT 1260
571 GluGluValAlaLeuGlnGluGlnValPheLysValProThrProIleThrGluMet 590
|||||
1261 GAGGAGAGTGTTCTGGAGAGCTCCAGGTTGTTAAGTGTCACACCTATCACCAGGAATG 1320
591 GluIleSerValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeu 610
|||||
1321 GAGATCTCTGTCAAAGGCAATGCTATACGTGGGCTCTGCGGTGGTGGTGGCGCCACTG 1380
611 ArgLeuHisGlnCysGluThrTyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAsp 630
1381 CGGTCGACCAATGATGAGATCTTACGGACATGCTGTGCGAGAGTGGCGCGCGGAGAC 1440
631 ProTyrCysAlaTyrAspGlyValaSerCysThrHisTyrArgProSerLeuGlyLysArg 650
1441 CCAATACGTGCTGGAGATGGTCTCTCTACCCACTACCGCCCGCCACTTGGCAGACGC 1500
651 ArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlnSer 670
1501 CGGTTCCGCGCGGAGACATCGGCGCAGCAGCAACCTCGCCCTGAGTCCGCGGCGCAAGC 1560
671 GlnGluGlnGluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGlnHisAsn 690
1561 CAGAGAGAGAGGAGAGTGGACTTGTGGCAGCCACCATGTCTACGGCAGGAGCAACAT 1620
691 SerThrPheLeuGlyCysLeuProLysSerProGlnAlaAlaValArgTyrPheLeuGln 710
1621 AGCAGCTTCTGGAGTGGCTCTCCCAAGTCTCCCGCAGCGTGTGGCGCTGCTTGGCAG 1680
711 ArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeuHisThrGln 720
1681 AGGCCAGGAGATGAGGGGCGCTGACAGGTGAAGAGCAGCAGCGAGCTTGGACACAGGAG 1740
731 ArgGlyLeuLeuPheArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrThr 750
1741 CGGGGCGCTGCTGTCCAGAGCTTACGCGTTTCATGGCGGAGCAGCTACACCTGACCACT 1800
751 LeuGlnHisGlyPheSerGlnThrValAlaArgLeuAlaLeuValIleValAlaSer 770
1801 CTGGAGCATGGCTTCTCCAGACATGTCGCTCGCTGTGGTGTGATTTGGCGCTCA 1860
771 GlnLeuAspAsnLeuPheProGluProLysProGluGluProProAlaArgGlyGly 790
1861 CAGCTGAGACACTGTTCCCTCGGAGCAGCAAGCAGAGGAGCCCGCGGGAGGCG 1920
791 LeuAlaSerThrProProLysAlaTyrTyrLysAspIleLeuGlnLeuIleGlyPheAla 810

```

```

Db 1921 CTGGCTTCACCCACCAAGCGCTGTACAGGACATCTCGAGTCAFTGGCTTCGCC 1980
Oy 811 AsnLeuProArgValAspGluTyrCysGluArgValIleTyrCysArgGlyThrThrGlyCys 830
Db 1981 AACCTGCCCGGGGTGATGATGATCTGTGAGCCGTGTGTGAGGGGCGCACCGGAATGC 2040
Oy 831 SerGlyCysPheArgSerArgSerArgGlyGlnAlaArgGlyLysSerThrAlaGly 850
Db 2041 TCAGGCTGCTTCGAGAGCGGAGCGGGGCAAGCAGGCGCAGAGCTGGCAGGG 2100
Oy 851 LeuGluLeuGlyLysLysMetLysSerArgValHisAlaGlnHisAsnArgThrProArg 870
Db 2101 CTGAGCTGAGGCAAGATGATGAGCCGGGTGATGCCAGCACAATCGACGCCCGG 2160
Oy 871 GluValGluAlaThr 875
Db 2161 GAGGTGAGGCGCAGC 2175

RESULT 7
AC094021
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 214046)
Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J., and Haugen, E.D.
Unpublished
Direct Submission
Submitted (13-SEP-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Sep 13, 2001 this sequence version replaced gi:14547325.

----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchigs@u.washington.edu
Drafting Center: BCM

----- Project Information
Center project name: chr-3
Center clone name: RP11-801L18 (bc0565)

----- Summary Statistics
Sequencing vector: M13; L08821; 50% of reads
Chemistry: Dye-terminator ET; 36% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Chemistry: Dye-terminator Big Dye; 53% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210653 bases at least Q40
Consensus quality: 212294 bases at least Q30
Insert size: 213746; sum-of-coverage
Quality coverage: 6.6x in Q20 bases; sum-of-coverage

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 14608: contig of 14608 bp in length
* 14609 14708: gap of unknown length

```



```

b 132497 ACCTCCTAGCAAGCCCGGAGGACCTACCTACAGATCTCAACTCTGGAGCCGATGTC 132556
y 131 ----- 131
b 132557 CCCAGACAGCTCCCGCTCTCANTCTGGGGGTGGGANTAGCAGACATCTGTGGGAGAAAA 132616
y 131 ----- 131
b 132617 TTGTCTTTGGCCGCATCCCTGAGCTGAGGTTGGGAGCCCTTCCCGCCATACCTTCCCTT 132676
y 131 ----- 131
b 132677 TTCTGACCCAGGAGCCATGTGCCAGGGGTCCCTCTGTGGAAACAGTTTCGCTGTGCC 132736
y 131 ----- 131
b 132737 CACCCTCAGCTGAGCCAGGCGCCCTGCAAAGACTCCCGCTTCTCCAGAGATCTTCCA 132796
y 131 ----- 131
b 132797 GCACCTTGTCTAGTTTCTGCTGTGGGGCCCTTCTTCCCTGGGACTGGTAAAGGACACA 132856
y 131 ----- 131
b 132857 GATGGAGAGTGGGAAGATGTATGCTGATGATGAAGGCTAAGGATAGAGGGTGGGA 132916
y 131 ----- 131
b 132917 GGTAGCCCCAGGGTAAAGGCTGGGCCAGGTGGAGGCTGTAGCCTCCGCTCCACCTGTC 132976
y 131 ----- 131
b 132977 CCCAGAGCCGTCGTGCCGAAGACCTTCCAGACACCATGTGATGGAACATTTTCCAGATG 133036
y 131 ----- 131
b 133037 TGAAGTGTACAGGGGTGATGCTACAGCCGGGGCCTCCACCTCCTTGGCTGGGGAA 133096
y 132 ----- 132
b 133097 CTAGGCTAGCTGTGGCCCTGATCTGCTCCCTCCCTTACATCTCTCTAGACCTCTGTCTGC 133156
y 1331 ----- 1331
b 133157 CAACCGCTCTGCCATCTTCTGGGCCCCCAGGGCTCCCTGAACCTCCAGGGCATGTACCT 133216
y 136 aasnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetLysIle 156
b 133217 AGATGAGTACCGAGACCGCTCTTCTGGGTGGCTGAGACGCCCTCTACTCTCTCGGCT 133276
y 176 uasPgluIyrArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgIle 176
b 133277 GGACCGAGCATGGCCAGATCCCGGGAGGT -GAGCTGAGCTGGCAGGGAAGAGGTCTCT 133335
y 186 ----- 186
b 133336 TTGAGAACTGGGCACTCCAGATGGAGGCGCAGATGGGAAACTGAGCCTGTGAGATCAG 133395
y 187 ----- 187
b 133396 GGGAGAGACTGAGTACGCGGGGCTGTGTTCCAGAGTCTGTGGCCACCGCAGGAGACA 133455
y 194 natrgluGluIcyValArgLysGlyArgAspProLeu----- 206
b 133456 GAGGAGAGAGTGTGTTCGAAGAAGGAGAGATCTTT -GTTGAGTGTCTGGGAGAGGATG 133514
y 206 ----- 206
b 133515 GTCCCGGACCTGACACCTTGCCCTGACCCCGCAGACAATCCGTGACACGCGCCAGGTA 133574
y 206 ----- 206

b 133575 CCTCTACAGGGGCAATTGGAGAGATCTTGAGGGCTTTTGAAGTCTGTGGGAATGTC 133634
y 206 ----- 206
b 133635 ACACCCACCCCTCCCGCAGAGCTGACCGCCAGACAGGGCTGAGCCTTGGCCCCAAGGA 133694
y 207 ----- 207
b 133695 GATGGTGGGACCCCTCTACCCCTCAACCCACCCCTCTCTTCTCCCGCAGACACGAGTG 133754
y 209 ----- 209
b 133754 salAsnPhenValArgValLeuGlnProHisAsnArgThrHisLeuAlaCysGlyThr 229
b 133755 CCGCAACTTCGTGCGGGTGTCTACAGCCTCACACCGGAGCCACCTGTCTGCTGTGGCAC 133814
y 229 rGlyAlaPheGlnProThrCysAlaLeuIleThrValGlyHisArgGlyGlu----- 246
b 133815 TGGGGCTTTCAGCCGACCATCGTGCCTCATCATCAGTTGGCCACCGTGGGGA -GTTGAGCC 133873
y 246 ----- 246
b 133874 TGGGCCAGGCCCCACAAGGAGCTGGAGGCGATGACATTTCTCTCCCTGCTGCTGACAT 133933
y 246 ----- 246
b 133934 TCCCATTTAGCCTGGGTAGGAGACAGAGGGCGCCTGAAGGCGCCTTGTGTGGG 133993
y 246 ----- 246
b 133994 GAGCTGACCCCATTTAGAGGGGCGCTGATGAGAGCGCTTCGCCCCACCCGCTCTCTG 134053
y 246 ----- 246
b 134054 GCGGCAAGGAGGCGTCTGCTGGGTGGGTGGCTGCTGCTCTCTCCAGGTGACGC 134113
y 247 ----- 247
b 134114 CCTCCCTCCCTCTCTCTCCCTCCACCTAGCATGTGCTCCTCAGCTGGAGCCTGAGCTGTGA 134173
y 257 userGlyArgGlyArgCysProHisGlyProSerArgProPheAlaSerThrPheIle-- 276
b 134174 AAGTGGCCGGGGGGGTGGCTCCTACAGAGCCAGCGTCCCTTGGCAGACCTTCAAT -AG 134232
y 276 ----- 276
b 134233 GNGGGTGAATGCCAGGCCAGGGGAGAGAAACAGGAGGGTGGCAGCTATGTCACTGG 134292
y 276 ----- 276
b 134293 GAAGGAGGGACCCGCTGCTGGGCTGGGCTCTCCCTGACTCTGTGCCCAACCCC 134352
y 277 ----- 277
b 134353 CACCCAGAGCGGGAGCTGTACACGGGTCTTCACTGTCTACTTCTGGGGCGAGAGGCCAT 134412
y 294 lIlePheArgSerGlyGlyProArgProAlaLeuArgSerAspSerAspGlnSerLeu 314
b 134413 GATCTTCGAGAGTGAAGTCTCTGGCCAGCTCTGCTTCCAGCTGTGACAGAGTCTCTT 134472
y 314 uHisAsp----- 316
b 134473 GCACGG -TGAGGCTGCCCCCTGGAAGTGGCAGGGGTGAAGAGGCTACTCTTCCAGCTAGG 134531
y 316 ----- 316
b 134532 ACACAGGCTCCAGTGGAGTCCCTGCCCCCATTTCTCTCTCCCACTGGGTCTACTCCA 134591
y 317 ----- 317
b 134592 TCTTCCCTGTGGGGCCCTCCACCCAGCATGACCTGTGACCTTCCAGAGACCCCGGTT 134651
y 319 eValMetAlaIleArgIleProGlnAsnSerAspGlnAspAsnAspLysValTyrPhePh 339
b 134652 TGTGATGGCGCGCGGATCCCTGAGAACTCTGACACGAGACAATGACAGAGTGTACTTCTT 134711

```

339 ephSerGluThrValProSerProAspGlySerAsnHisValThrValSerArgVa 359
|||||
Db 134712 CTTCTGGAGAGGCTCCCTCGCCCGGATGGTCTGAAACATGCTACAGCCCGCT 134771
359 lGlyArgVal 362
|||||
Db 134772 GGGCCGGCTCTGCGGTGAAGAGCTGGGAAGGGGTGTGAGTTCAAGTCCCTGCTACAC 134831
362 362
Db 134832 CCTCCATTCCAGACCTTTCTCCCCCAACCCCTGGCCAGACAGTAGGAAATGAGAGCAA 134891
362 362
Db 134892 GGATGCAAAAGCTGACATGATACCTCTCCCGAGTGGGTGAGGTAGAAAGTGGATGCT 134951
363 367
CysValAsnAspAl
Db 134952 GGGACCCATCTCTGCGGGGCCCAAGCGGTGACTCTGGCCCTTATCTCCAG -AATGATGC 135010
367 387
AGlyGlyGlnArgValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCy 387
|||||
Db 135011 TGGGGCCAGCGGGGTGCTGTGAACAATGAGACACTTCTCAAGGCCAGGCTGCTG 135070
387 402
sSerValProGlyProGlyGlyAlaGluThrHisPheAspGlnLeu 402
|||||
Db 135071 CTCGGTCCCGCCCTGCTGTGTGTCGAGACCCACTTTGACAGCT -AGTGGAGGCTGG 135129
402 402
Db 135130 CCAAGCAGGCTAGAGGCTCGGAGTGGGAGGCTGAGACTGACAGGCAATTGG 135189
403 410
GluAspValPheLeuLeuTrpPr 410
|||||
Db 135190 CGGTGGCATGACATCATACCTTTCTCTACCCAGAGAGTGTCTCTGCTGTG 135249
410 425
OLysAlaGlyLysSerLeuGluValTyraLeuPheSerThrVal 425
|||||
Db 135250 CAAGCCGCGGAAGACCTGTGAGGTGTACGCGCTTCACGACCTGAGGTGGCACACCC 135309
425 425
Db 135310 AGCACCCGCGCTGGAGTGGGCCCTGTCTTGTCCGATTCGTGACCTGTGACTGT 135369
425 425
Db 135370 ACTCTGACTCTCTCTTGGCTTTGGGCTTTGCTTTGGAGTCTGCTG 135429
425 425
Db 135430 CTGAGGACACTGACACCTCTGTCTGCTGACATGGAACAGGCTGTCCCTTCTGTCTT 135489
426 433
SerAlaValPheGlnGlyPheA 433
|||||
Db 135490 GATGGCACACTGCGCTGACTGTCCCATGTGCCCCGCGAGTCCGCTTCAGGGCTTCG 135549
433 458
laValCysValTyraHisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisA 458
|||||
Db 135550 CCGTCTGTGTACCAATGCGACATGTGGAGGTTCATCAAGGGGCCCTTTGGCCAC 135609
453 473
rGAspGlyProGlnHisGlnTrpGlyProTyraGlyLysValProPheProArgProG 473
|||||
Db 135610 GAGATGGGCTCAGCACAGTGGGGCTTATGGGGCAAGGTGCCCTCCCTCGCCCTG 135669
473 474
lyVal 474
Db 135670 GCGT -GGTGAATATCTGCTGGGCCAGGCCAGGAGAGGCTGAGCAGATGTGGGA 135728
474 474
Db 135729 AGGGCTCTTGGAGTGGCTCATGTGTCCCATGACCCCAAGGAGGAGGAGCTTGGCCA 135788

474 474
Db 135789 GCCTGGGCTCCAGAGGGCTGGGGTGGGTGGGTGTAGTGAAGAGAGACCCCTC 135848
474 474
Db 135849 CTGCACTGAGAGAGTCCCTGTTGGGTGAACAAGACTGGCGGTGAAGGTGTGTG 135908
475 480
CysProSerLysMet 480
Db 135909 CTGTCTCAGATCTTACAGCACTCCCTCTCCCTGCTGCCCCACAGAGGCCCGCAAGATGA 135968
480 500
hrAlaGlnProGlyArgProPheGlySerThrLysAspTyraProAspGluValLeuGlnP 500
|||||
Db 135969 CCGCACAGCCAGAGAGGCTTTTGGCAGCACCAAGGACTACCAAGATGAGTGTGCTGACT 136028
500 520
heAlaArgAlaHisProLeuMetPheTrpProValArgProArgHisGlyArgProValL 520
|||||
Db 136029 TTGCCCCAGGCCACCCCTCATGTCTGTGCTGTGCGCTGCGACATGGCCGCTGTCC 136088
520 540
euValLysThrHisLeuAlaGlnGlnLeuHisGlnIleValValAspArgValGluAlaG 540
|||||
Db 136089 TTGTCAAGACCCACCTGCGCCAGCAGCTACACAGATGTGTGACCGCTGAGAGCG 136148
540 551
luAspGlyThrTyraSpValIlePheLeuGlyThr 551
|||||
Db 136149 AGGATGGAGCTTACGATGCTATTTCTGCGGAC -TGATAGACAGCTGAGCCAGGCTGG 136207
551 551
Db 136208 CCAAGCTGAGAGGGGTGTGGGTCTGTGCTGTCCCTGCTGCCCCGTACCCCTCT 136267
552 568
AspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlyLysSera 568
|||||
Db 136268 CCCTGGCCCCAGACTAGGGCTGTGTGTCAAAATGATGTGCTCCAGCAGGGGCTCAG 136327
568 582
laGluProGluGluValValLeuGluGluLeuGlnValPheLys 582
|||||
Db 136328 CTGAACCTGAGGAAGTGTGTGTGAGGAGCTCCAGGTGTTAAAGTTAAAGCAAGGGGCTGA 136387
582 582
Db 136388 TGAAGGGTCCCTCCAAAGCCCATCTGTGGTCTGAGCCTGGCTACAGGATGCAAGAGC 136447
582 582
Db 136448 AAGCCAGTAGGAGCCAGCCTGTCTTACTGAATTTGACAGCTGGCAGAGTGTGGGAG 136507
582 582
Db 136508 GGGATGCAAGAAAGATGCCCCCTCCATCCAGCCGACCTCGCTGTGGCAGAGGCTCGC 136567
582 582
Db 136568 AGCCACTTATGATGACTGAGAGGCCAGAGTACAGCTGTGAGCTGAAGTCCGTGGGG 136627
582 582
Db 136628 TGCTGGAGGATGTGCAAGGGAACAACCTTCTGCAAAACCCCTGAGAGCTCAAGTTGC 136687
582 582
Db 136688 AAGGAGAGATGTTGAGATCCCAAGGCCCTTGAATGTCTCTTGGCCCTCTAGAGACT 136747
582 582
Db 136748 CAGGAAGGGGTCACTGTGTGTTTGGGGGCTTGTGCTCTGAGACCTGAGCTGGGTA 136807
582 582
Db 136808 TTGTGACAGCTGCTCATTTGCTGTGGGAACAAGCTGAGAGAGGAGACACTGTAGCTGT 136867
582 582

Db 136868 GCTTCAGTCAAGCTCTGCTGTGGAGGCCAGCATGACAGGGGCTTCTGTAAAC 136927
OY 582 ----- 582
Db 136928 CATGCTGTGGAAAGACTGTCTAGCCTTAGCTGCCAGGAGTGCAGACCCAGCCTGGGA 136987
OY 582 ----- 582
Db 136988 GCTAGGTGTTCCACCACCGGTTGGAGAGACGTGAAGCAGGGGTGCAGGGCAGGTC 137047
OY 583 -----ValProThrProIleThr 588
Db 137048 TTGGTCAACCCGACCCACACTGAGACCTTCTCTCCAGGTGCCAACCCTATCAC 137107
OY 589 GluMetGluIleSerValIysArg ----- 596
Db 137108 GAATGTGAGATCTCTGTCAAAAGGTAAAGACAGCTCCACCTCTCTTGAGCCCCCT 137167
OY 596 ----- 596
Db 137168 TCTGCCCTGTCCGAGGCTCTCTCTGCCCCCTGCTCCGACGCTCTGAGCTCCAC 137227
OY 596 ----- 596
Db 137228 TGTAGCTGTGGGAAAGCTGTCTGAGGCGAGGCTGATTCGCCCTGCAACAAGTGGCCAG 137287
OY 596 ----- 596
Db 137288 TGGGCACTGGAGGCAACCGAGTAGTGGGGCCATGTCAGATTAAGGGATAGTCACAAAG 137347
OY 596 ----- 596
Db 137348 CCAGTTCCTAAGGAGGCAAAAGTGTGCTGAAAAACAAATCCAGATGTTAGTCCAA 137407
OY 596 ----- 596
Db 137408 CAAAGCCACTTTTAGAGGCTACTATATTGGCCAGTGGCAAAACCCTGCCCATAT 137467
OY 596 ----- 596
Db 137468 TTATCTGCTCCAGTGTATCTTTGTAAATCCTAACTTTCGTGCAAGGTTGCTTAAAGC 137527
OY 596 ----- 596
Db 137528 AAGTGCACCTTAAGTGCAGGACAACAGATAAAGCCCTGGGATGAGGGCAAA 137587
OY 596 ----- 596
Db 137588 GGTCAACACCTGGAGGGCTGGAGAAGCCCTTAGAAGAAACATCCAGGCTTGGTTC 137647
OY 596 ----- 596
Db 137648 TCCATCATCCCTTGTGTCTGTGCACAGAAATCTGGTATGCAGAGATGGGTGATAACA 137707
OY 596 ----- 596
Db 137708 CTTTGTGTTTTTAATGAATCAAGAGGTTCTGGAAGACTTCTGCAAGTGGCAGGGGT 137767
OY 596 ----- 596
Db 137768 CTCAGGAAGCTGAGGGGCTGAGAGCCTGGCTGTGGTGGGCGAAGCCTGGGCTCACTCA 137827
OY 597 -----GluMetLeuTyrValIleSerArgLeuGlyValAla 608
Db 137828 GGTCTGTGCTGTGCTCCCTCCGCAAAATGCTATACGTGGGCTCTCGGCTGGGTGGCC 137887
OY 609 GluLeuArgLeuHisGluCysGluIleThrTyrGlyThrAlaGlyAlaGluCysCysLeuAla 628
Db 137888 CAGTGGCGGCTGTCCAAATGTAGACTTAAGGCACTGTGTGAGAGTGTGCTGGCC 137947
OY 629 ArgAspProTyrCysAlaIlePaspGlyAlaSerCysThrHisTyrArgProSerLeuGly 648
|||||

Db 137948 CGGGACCCATACTGTGCTGGGATGTGCTCTCTGTACCACCTACGCCACCCAGCTTGGC 138007
OY 649 LysArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuIleuGly 668
Db 138008 AAGCGCGGTTCCCGCGAGAGATCCGGACGGGCAACCTGCTGTGAGTGGGCG 138067
OY 669 GlnSerGlnGlu ----- 672
Db 138068 CAGAGCCAGGAGGTGAGTGTGGCAGGTGGGAGGGCCACAGACTGGGCTGGAATCATCC 138127
OY 672 ----- 672
Db 138128 CTGTGCCCGGCTCCACCCCTGCGAAGTGCATTTCAAGATTGTCTTGCATAGCAGATA 138187
OY 672 ----- 672
Db 138188 GTCTCAAGTTTCCCTCTCTGTGTAGCTCTCTGTAAGTGAAGCAGGCTTTGGACA 138247
OY 672 ----- 672
Db 138248 CAGTGGCTGGGCTGGGATCCACGCAATGCTTAAGTCTTGTGATACTAGTGGGGCTCTG 138307
OY 672 ----- 672
Db 138308 AGCTGTGTGTCAAGTGTGACCTGGAGTCTGGTCCCATAGTTCACAGCTGAGCCAA 138367
OY 673 -----GluGlnAlaValAlaGlyLeuValAlaIleThrMetValTyrGlyTh 687
Db 138368 CCTCTGCTCTCTCCAGAGAGGAGTGGACTTGTGGCAGCCACCATGTCTACGGCAC 138427
OY 687 rGluHisAsnSerThrPheLeuGluCysLeuProLysSerProGlnAlaAlaValArgTr 707
Db 138428 GGAGCAACAAATAGCACCTTCCGAGTGGCTGCCCAAGCTCCCGAGGCTGTGGGCTG 138487
OY 707 PheLeuGlnArgProGlyAspGlyProAsp ----- 718
Db 138488 GCTCTTGCAGAGGCCAGGAGGTGAGGGGCTGACCAAGTGAAGTGAAGAGACCCCTGCTTC 138547
OY 718 ----- 718
Db 138548 TAGCCTGTGTGTCTGAGGTCACTCCATCCGAAACCCCAAGACCTGAGCCCGAGA 138607
OY 718 ----- 718
Db 138608 AGGGGTTCTGGGGGAATCTGCCAAGATTAACTTAAGTTTGTATTAGCTCTGG 138667
OY 718 ----- 718
Db 138668 GGTGCTAGAAAGAGTGCCTATTTTTGCCTTTTGGAAATGCCCAAGCCAGATGTGCT 138727
OY 718 ----- 718
Db 138728 GGAACATCCCCAGTCTGAGGGGTGGTGTATGAGAGAGAGAGAGAGAGAGGCTGGGAGGCC 138787
OY 718 ----- 718
Db 138788 AGTAAGTCTCCATGTTTAAACACATGGCTCATCTTTGAAGATTCTCTTCCAGGGGT 138847
OY 718 ----- 718
Db 138848 TCTTGGTTCCTTCTGTGCCCAAGAGGAGGGGAGGGGCTGTGCCCTGGCCAGAGCTAGGC 138907
OY 718 ----- 718
Db 138908 TTCAAGAACTCCATATGCTACACACACACAAGAGAGAACTTGTGGTCCCTAAACA 138967
OY 718 ----- 718
Db 138968 CTGGTCTTGCCTCCACCTCAAAAGCTGGCCCCACACTTCTCTGGAATCCAGCTG 139027
OY 718 ----- 718
Db 139028 TCTGCTCAGGGGCTCTCTCCCTCCCATCCGGAAGGTCTATCAACCCCAAGATGTTTAA 139087

sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of bases' quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

Location/Qualifiers

```

FEATURES
source      1..123943
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="3p21.1-9"
             /clone="RPC14-793p23"
             2078..2357
             /rpt_family="AluX"
             /complement(4381..4674)
             /rpt_family="AluSg"
             /complement(6668..6724)
             /rpt_family="MIR"
             /complement(6725..7029)
             /rpt_family="AluSc"
             /complement(7040..7197)
             /rpt_family="MIR"
             /join(7498..7643,7987..8221,8767..8939)
             /gene="OU12F08.X1 Homo sapiens CDNA A1004779"
             8793..8921
             /standard_name="G27739"
             /db_xref="dbSTS:41594"
             9133..9283
             /standard_name="D3S3975"
             /db_xref="dbSTS:10759"
             /complement(join(9446..10863,11044..11116,11221..11313,
             1180..11740,11858..12336,12895..13028,13552..13736,
             14204..14351,14692..14815,15268..15346,15613..15755,
             15838..15899,16397..16516,16913..17045,17993..18047,
             18153..18183,18281..18320))
             /gene="Human mRNA for KIAA0272 gene D87462"
             9497..9735
             /standard_name="D3S3829"
             /db_xref="GDB:626512"
             13712..13962
             /function="Low coverage"
             /complement(18321..18407)
             /rpt_family="GC-rich"
             18522..18742
             /function="Low coverage"
             /complement(18554..18661)

gene
             /rpt_family="GC-rich"
             /join(19128..19686,21251..21360,22436..22488,22936..23027,
             28273..28374,28751..28875,29343..29502,30088..30166)
             /gene="Unigene cluster containing AJ223811, AA280976, and
             A1199338"
             20373..20654
             /rpt_family="AluJb"
             20931..21138
             /rpt_family="MIR"
             21525..21781
             /rpt_family="MIR"
             21993..22279
             /rpt_family="AluSx"
             23167..23448
             /rpt_family="AluJb"
             /complement(23639..23858)
             /rpt_family="AluSg"
             /complement(23859..24621)
             /rpt_family="L1ME2"
             /complement(24622..24923)
             /rpt_family="AluY"
             24932..25002
             /function="Low coverage"
             24993..25290
             /rpt_family="AluY"
             /complement(25299..25521)
             /rpt_family="L1ME2"
             /complement(25497..26685)
             /rpt_family="L1ME2"
             27118..27246
             /rpt_family="MIR"
             /complement(27249..27544)
             /rpt_family="AluSx"
             27698..27826
             /rpt_family="MIR"
             /complement(27868..28170)
             /rpt_family="AluY"
             31632..32081
             /note="Region: q987a10.x1 Homo sapiens CDNA A1221496"
             31774..31924
             /standard_name="G38576"
             /db_xref="dbSTS:58631"
             31932..32081
             /standard_name="G30450"
             /db_xref="dbSTS:44631"
             32599..32898
             /rpt_family="AluSg"
             33780..33819
             /rpt_family="(CA)n"
             /complement(33968..33996)
             /rpt_family="AT-rich"
             /complement(34011..34082)
             /rpt_family="AluS"
             /complement(34083..34121)
             /rpt_family="(GAAA)n"
             /complement(34123..34410)
             /rpt_family="AluSg"
             /complement(34413..34697)
             /rpt_family="AluJb"
             /complement(34718..35265)
             /rpt_family="L1MC3"
             /complement(35266..35558)
             /rpt_family="AluSg"
             /complement(35561..35738)
             /rpt_family="L1MD3"
             /complement(35739..36036)
             /rpt_family="AluSg"
             /complement(36038..36236)
             /rpt_family="L1MD3"
             /complement(39127..39419)
             /rpt_family="AluSx"
             41664..42236
             /note="Region: Unigene cluster containing W95768 and

```


131 ----- 131
51766 GCCTGTGGCCACCCTCCAGCTGAGCCAGCGGCGCTGCAMAAGATCCGCGCTTCCTC 51707
131 ----- 131
51706 CAGGATCTTCCAGCACTTGTCTACTTCTGCTGTTGGGGCCTTCTCCCTGGGACT 51647
131 ----- 131
51646 GGGTAGGGCAGATGGAGCACTGGGAGATGTTAGCTGATGATGAAGGCTAAGAAAT 51587
131 ----- 131
51586 GGAGGGTGGAGTAGGCCCCAGGCTAGCGTGGCCAGGTGGAGGCTTGAGCCTCCC 51527
131 ----- 131
51526 GTCCCACTGTCCCAAGAGCCATGTCGAAAGCCTTCCAGCACCATGTGGATGGAAC 51467
131 ----- 131
51466 ATTTTCAGATGATGTGTGTCAGGGGTGATGCTCAGCGGGGCTCCACCTCCACTCC 51407
132 ----- 132
51406 TTGGCTGGGGAAGTGAAGCTGAGCTGTGGCCCTGATCTGCTCCCATACATCCCTTGA 51347
132 PLeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGly 152
51346 CCTCTGTCTGGCAACCGCTCTCCATCTTCTGGGGCCCGCCAGGGCTCCGAACTCCA 51287
152 nAlaMetIleuAspGluTyrArgAspArgIleuPheLeuGlyIleuAspAlaLeuTyr 172
51286 GGCCATGTACCTAGATGATGATACCGAAGCCCTCTTCTGGGTGGCTGAGCGCTCTCA 51227
172 rSerLeuArgIleuAspGlnAlaIlePheProAspProArgGluVal ----- 186
51226 CTCTGTGGGCTGGACAGGCAATGGCCAGATCCCGGGAAGT -GAGCTGAGCTGGCAGGG 51188
186 ----- 186
51167 AAGAGGTCTGTTTGAAGACTGGGCACTCCAGATGGGAGGCCAGATGGGAAACTGAGC 51108
187 ----- 187
51107 CTGTGATGATCAGGGGAGGAGCACTGAGGTCAAGCGGGCTGTCTCCAGGTCTGTGGCCACC 51048
190 oGlnProGlyIleuArgGluGlyValArgIleuAspProLeu ----- 206
51047 GCAGCCAGGACAGAGGAGAGTGTGTGAAAGGAGAGATCTTT -GSTGAGTGTCTG 50989
206 ----- 206
50988 TCGGGAGGATGTCCTCCCGACCTGACACCTGCTCCCTGACCCCGAGAGCAATCCGTGAC 50929
206 ----- 206
50928 ACGCCCAAGTAACCTCTTACAGGGGCAATTGGAGCAGATCTTGAGGGCTTTTGAAGTCT 50869
206 ----- 206
50868 GCTGGGATGTACACCCACCCCTCCCGAGCTGACCGCCAGACAGGGCTAGACCTT 50809
206 ----- 206
50808 GGGCCCAAGAGATGTTGGGACACCCCTCTCACCCCTCAACCAACCCCTCTTCTCCCC 50749
207 ----- 207
50748 GCAGACAGAGTGGCCACTTGTGGGGGTGTACAGCTTACACAGCGGAGCCAGCTGCT 50689
225- uAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuIleThrValGlyHisArgGly 245

|||||
Db 50688 AGCCTGTGGACTGGGGGCTTCCAGCCACCTGTGCCCCATACAGTGTGGCCACGCTGG 50629
245 yGlu ----- 246
Db 50628 GGA -GGTAGCCTGGGCCAGGCCCCACAGAGGACTGGAGGCAATGACATTTCTCTCTG 50570
246 ----- 246
Db 50569 CTGCTGATCATTCATTTAGCCTGGGTAGGAGAGAGGGGAGGGGCTGAAAGGGCG 50510
246 ----- 246
Db 50509 CTTTGTGTGGGAGCTGACCCCATTTAGAGGGGCCCTGATGAGAGACGCTTCCGCCAC 50450
246 ----- 246
Db 50449 CCGCGCTCTGGCGGACGAGGAGGCTCTGCTTGGGTGGGCTGCTGCTTCTCT 50390
247 ----- 247
Db 50389 CCCAGTGAGCCCTCCCTCTCTCTCTCCCTCCACTAGCATGTGCTCCACTGGAGCC 50330
253 oGlySerValGluSerGlyArgGlyArgCysProHisGluProSerArgProPheIle 273
50329 TGGCAGTGTGGAAGTGGCGGGGGGCTGCTTACAGACCCAGCCCTTCTTGGCAG 50270
273 rThrPheIle ----- 276
Db 50269 CACTTTCAT -AGTGGGTGATGCCAGGCCAGGAGGAGAACAGAGGATGGCAGCTGGC 50211
276 ----- 276
Db 50210 CTATGTCACTGGAGAGGAGGAGCCGCGCTGCTGGGCTGGGCTCTCCCTGACTGTG 50151
277 ----- 277
Db 50150 TTCCCCACCCCCACCCAGAGGAGGAGCTTACAGGAGTCTCACTGCTGACTTCTGG 50091
290 yArgGluAlaMetIlePheArgSerGlyIleProArgProAlaLeuArgSerAspSer 310
50090 GCGAGAGGCGCATGATCTTCCGAATGGAGGTCTGTGGCAGCTCTGCGTTCGACTCTGA 50031
310 pGlnSerLeuLeuHisAsp ----- 316
Db 50030 CCAGAGTCTCTTGCACGG -TGAGGCTTCCCTGTGAACTGGCAGGGGTGAAAGGCTTACC 49972
316 ----- 316
Db 49971 TTCCCACTAGGACACCAAGCTCCAGTGGAGTCCCTTGCCTCATTTCTCTCCCACTGG 49912
316 ----- 316
Db 49911 GTTCTACTCCATCCCTCCCTGTGGGGCCCTCCACCCAGATGACCTGTACCCCTTCC 49852
317 ----- 317
Db 49851 AGACCCCGGTTGTGTGGCGGGCCGATCCCTGAGAACTGTGACAGAGCAATGACAA 49792
335 sValTyrPhePheSerGluThrValProSerProAspGlyIleSerAsnHisValh 355
49791 GGTGTACTTCTTCTTCTCGAGAGAGGTCCCTCGCCCATGTGGCTGCAACATGTCTAC 49732
355 rValSerArgValGlyArgVal ----- 362
Db 49731 TGTACAGCGCGGTGGCGCGTCTGCTGTAAGAGCTGGAAAGGGGTGTGAGTTCAAG 49672
362 ----- 362
Db 49671 TCCCTGTACACCTCCCATTTCCAGACCTTTCTCCCCCAACCCCTGGCCAGCAGTAGGG 49612
362 ----- 362

Db 45235 AGATGGTGTTCACACACTAGCTAATGGTCAAGCAATCCACCCACACACAGATG 45176
OY 718 ----- 718
Db 45175 TGCAGACTAAACCCACACAGAGGTCCAGGCCAGGCGGTGGATCAGGGCGGCACTC 45116
OY 718 ----- 718
Db 45115 CCACCTCCACGCTGTATGGCGATCAGCCACCCTGGGCTGGGCTGCAGCTGCAC 45056
OY 718 ----- 718
Db 45055 CGTTGGGGATGAAGGCTTCAGACACAGATGGAAAGCCCTGTGTGTCTGTGTGT 44996
OY 718 ----- 718
Db 44995 GTGTGGCTGACGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 44936
OY 718 ----- 718
Db 44935 AGTTTCTGAAGCCAGTTCCTGGGCGACAGGCGCACAGCCCTCCCTTGAGAAAGCCTGTA 44876
OY 718 ----- 718
Db 44875 GATGGATGATGACTCCAGAGAGGCTAAGGGATACATCTTACCCCTATGTTCTTGG 44816
OY 718 ----- 718
Db 44815 GAATGTGTCCAGGCGCTCTGAGGCCAAGATTGAGTCTCTGATTAAGGAGAAACCA 44756
OY 718 ----- 718
Db 44755 GCCAGTCAATCTTGAAGAGGCGCTCTGTGGCTATGGAGGTGAGGCGCCACCTGA 44696
OY 718 ----- 718
Db 44695 ATTAGT 44636
OY 718 ----- 718
Db 44635 GGAGATAGGGCAGAAAGTGGCCAGGCACTGGGCCCAAGATGGCAGAAAGCTGGGGC 44576
OY 718 ----- 718
Db 44575 AGAGAGCCGATGTAACTCCACCTTGCACCTACTACGCGCTCCCGCCGCGATGGC 44516
OY 718 ----- 718
Db 44515 AGGTGAAGAGGAGCGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 44456
OY 718 ----- 718
Db 44455 AGGTGAAGAGGAGCGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 44396
OY 718 ----- 718
Db 44395 GGT 44336
OY 718 ----- 718
Db 44335 AGCCAAAGCCAGAGGAGCGCGCCAGCCCGGAGAGGCTGTGTGTGTGTGTGTGTGT 44276
OY 718 ----- 718
Db 44275 GGT 44216
OY 718 ----- 718
Db 44215 GTGAGCGGT 44156
OY 718 ----- 718
Db 44155 GGGGCAAGAGCGCGAGGCGCGCAAGAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGT 44096

OY 859 erArgValHisalagIuHisasArgrProrArgIuValGluAlaThr 875
Db 44095 GCCGGGTGATGCCAGACATCGAGCGCCCGGAGAGTGGAGCCACG 44046
RESULT 9
AX155172 2337 bp DNA linear PAT 22-JUN-2001
LOCUS sequence 3 from Patent WO0140278.
DEFINITION
ACCESSION AX155172
VERSION AX155172.1 GI:14536657
KEYWORDS
ORGANISM
SOURCE
synthetic construct.
artificial sequence
1 (bases 1 to 2337)
REFERENCE
Holloway, J.L. and Foley, K.P.
Human semaphorin zsmf-16
Patent: WO 0140278-A 3 07-JUN-2001;
TITLE
JOURNAL
Zymogenetics, Inc. (US)
location/Qualifiers
FEATURES
source
1. 2337
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Degenerate Polynucleotide sequence of zsmf-16"
BASE COUNT 308 a 273 c 468 g 293 t 995 others
ORIGIN
Alignment Scores: 2337
Pred. No.: 3102.50
Score: 71.75%
Percent Similarity: 71.75%
Best Local Similarity: 65.37%
Query Match: 6
Gaps: 2
US-09-813-290-2 (1-875) x AX155172 (1-2337)
OY 94 MetaIarProserAlarPalalIecySTRPLeuLeuGlyLeuLeuHisGlyGly 113
Db 1 ATGCNCNCNMSGNCNTGGCNAHTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
OY 114 SerSerGlyProserProGlyProSerValProArgLeuArgLeuSerTyArgAspLeu 133
Db 61 WSNWSNGNCNMSGNCNCGNCNMSGNCNMSGNCNMSGNCNMSGNCNMSGNCNMSGNCNMS 120
OY 134 Leu-----Ser 135
Db 121 ATGCTMGNARCNMSNMSNACNATGTGATGATGATGATGATGATGATGATGATGATGATGAT 180
OY 136 AlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyT 155
Db 181 GCNAAYGNMSNCGNATHTTGT 240
OY 156 LeuAspGlyArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeuTySerLeuArg 175
Db 241 YTNAGVARGATMNGNAYMNGVNTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
OY 176 LeuAspGlnAlaTrpProAspProArgGluValLeuTrpProGlnProGlnGlyAlaTyr 195
Db 301 YTNAGVARGATMNGNAYMNGVNTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 330
OY 196 GluGluCysValAlaGlyGlyArgAspProLeuTrpGlyCysAlaAsnPheValArgVal 215
Db 331 -----ACNGARTGVCMAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 357
OY 216 LeuGlnProHisAsnArgTrpHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThr 235
Db 358 YTNARCCNCAIAYMKNACNCAIYNTYNTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 417
OY 236 CysAlaLeuIleThrValGlyHisArgGlyGlnHisValLeuHisLeuGlnLeuProGlySer 255
Db 418 TGYCCTNTNATHACGNTGNGCAYMNGNARGARCAVGTNTNCAIYTTGARGCCNGGMSN 477

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT

Coyle, M.D., Dathorne, S.R., David, R., Davila, M., Davis, C.,
Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Drepper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Fells, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Guarinate, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, S., Huber, J., Hulyk, S., Hume, J.,
Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudan, C., Karlsson, E., Kelly, S., Khan, U., King, L., Kotvch, J.,
Kovar, C., Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W.,
Louisged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R.,
Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
Matthez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabati, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Ogaunye, N., Oviado, R., Pace, A., Payton, B.,
Ogund, A., Okmondu, G., Orangun, N., Nickerson, E., Nwokenkwo, S.,
Peeney, J., Petruz, L., Peters, L., Pickens, R., Plinius, E., Pu, L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Severy, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N.,
Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabot, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, C., Taylor, T.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 229706)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901619.

Genome Center

```

** NOTE: Estimated insert size may differ from sequence length.
** (see http://www.hpsc.bcm.tmc.edu/docs/genbank/draft\_data.html)
** NOTE: This is a "working draft" sequence. It currently
** consists of 32 contigs. The true order of the pieces
** is not known and their order in this sequence record is
** arbitrary. Gaps between the contigs are represented as
** runs of N, but the exact sizes of the gaps are unknown.
** This record will be updated with the finished sequence
** as soon as it is available and the accession number will
** be preserved.
*
* 23933: contig of 23933 bp in length
* 1
* 24033: gap of unknown length
* 23934
* 45428: contig of 2195 bp in length
* 24034
* 45528: gap of unknown length
* 45429

```

45529	59289:	contig of 13761 bp in length
59290	59389:	gap of unknown length
59390	73166:	contig of 13777 bp in length
73167	73266:	gap of unknown length
73267	86647:	contig of 13381 bp in length
86648	86747:	gap of unknown length
86748	100569:	contig of 13822 bp in length
100570	100669:	gap of unknown length
100670	115083:	contig of 14414 bp in length
115084	115183:	gap of unknown length
127043	127042:	contig of 11859 bp in length
127143	127142:	gap of unknown length
134393	134393:	contig of 7251 bp in length
134494	134493:	gap of unknown length
142907	142906:	contig of 8413 bp in length
143007	142951:	gap of unknown length
143007	142951:	contig of 5245 bp in length
148352	148351:	gap of unknown length
148352	154125:	contig of 5774 bp in length
154126	154225:	gap of unknown length
154226	162069:	contig of 7844 bp in length
162070	162169:	gap of unknown length
162170	162180:	contig of 5111 bp in length
167281	167380:	gap of unknown length
167381	174948:	contig of 7568 bp in length
174949	175048:	gap of unknown length
175049	179089:	contig of 4041 bp in length
179090	179189:	gap of unknown length
179190	183664:	contig of 4495 bp in length
183685	183784:	gap of unknown length
183785	189316:	contig of 5533 bp in length
189317	189416:	gap of unknown length
189417	195866:	contig of 6450 bp in length
195867	195866:	gap of unknown length
195967	201123:	contig of 5156 bp in length
201123	201222:	gap of unknown length
201223	205343:	contig of 4121 bp in length
205344	205443:	gap of unknown length
205444	208257:	contig of 2814 bp in length
208258	208357:	gap of unknown length
208358	211417:	contig of 3060 bp in length
211418	211517:	gap of unknown length
211518	214163:	contig of 2646 bp in length
214164	214263:	gap of unknown length
214264	217312:	contig of 3049 bp in length
217313	217312:	gap of unknown length
217413	219666:	contig of 2554 bp in length
219667	220066:	gap of unknown length
220067	221858:	contig of 1792 bp in length
221859	221958:	gap of unknown length
221959	224415:	contig of 2457 bp in length
224416	224515:	gap of unknown length
224516	225533:	contig of 1018 bp in length
225534	225633:	gap of unknown length
225634	226937:	contig of 1364 bp in length
226938	227097:	gap of unknown length
227098	228527:	contig of 1430 bp in length
228528	228627:	gap of unknown length
228628	229706:	contig of 1079 bp in length.

Alignment Scores:	1.33e-81	Length:	229708
Pred. No.:	2287.50	Matches:	648
Score:	26.19%	Conservative:	31
Percent Similarity:	24.99%	Mismatches:	64
Best Local Similarity:	48.09%	Indels:	1854
Query Match:			

```

DB: 2 Gaps: 16
US-09-813-290-2 (1-875) x AC095672 (1-229706)
OY 129 SerTyrArgSpLeuLeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlyGlySer 148
Db 91625 TCCTCTGATGATCTCTGATACCAACCCGCTCTGCCATCTTTCTGGGTCCAGGGCTCC 91684
OY 149 LeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArgAspArgLeuPheLeuGlyLeu 168
Db 91685 CTAGACCTCCAGATCATATACCTGGATGATGATGCGGACCGCTCTTTCTGGAGGCCGA 91744
OY 169 AspAlaLeuTyrSerLeuArgLeuAspGlnAlaTyrProAspProArgGlyVal 186
Db 91745 GATGCCCTTACTCTCTGTGAGACTGTGATCAGCATGCGCCAGATCCCGGAGGT-GAGCCG 91803
OY 186 -----
Db 91804 GACTAAGGTGGTGAAGAATAGGACCTTTCAGGTGAGAAAGATGGCGAGACTTGAAGACA 91863
OY 187 -----
Db 91864 TGATTTAAGACCTTGAATGCTTGTGCTGTGCTGTTCAGGTCTTGTGCGCCGCGAG 91923
OY 192 ProGlyLysArgGluGlyLysValArgLysGlyArgProLeu 206
Db 91924 CCGGACAGAGAGTGAATGTCTCCGAAAGGAAAGACCTTTGGTGTGATTTGTTGGC 91983
OY 206 -----
Db 91984 AAGGATACCCCTCCCGCGGTTCTCACAAATGTCTGACGACACTCCCGTGTGCTCTGC 92043
OY 206 -----
Db 92044 TGGAGAGATTGGAGCAGGACACTGTCTTCAGGACATCTGGAGAGATGTGACCTTCACC 92103
OY 206 -----
Db 92104 GTTCTCAGAACTGACACTGACACAGAGAGCTAAGGAGAAAGCTTAGAACCCAG 92163
OY 207 -----
Db 92164 GCTACCCCTCAGCAGCCCATTTCTTCCCTACAGACCGAGTGTCCCACTTCTGCGGGT 92223
OY 215 LLeuGlnProHisAsnArgTyrHisLeuLeuAlaGlySerGlyThrGlyAlaPheGlnPro 235
Db 92224 CTTCACACCCCAACAACGAGCCCAATCTGTGCTGCGGCACTGCGCTTCCAGCCAT 92283
OY 235 rCysAlaLeuIleThrValGlyHisArgGlyGlu 246
Db 92284 CTGCACCTTCACTACAGTGGGGCATCGGGGGA-GGTGAGCTGTGGTGTGCCCATGAG 92342
OY 246 -----
Db 92343 AAGTTGGAGAGTGTGACATTTCTCCCGCTGTGACACATTTCCCATTTGAGCAGAGGT 92402
OY 246 -----
Db 92403 AAGGCGAGAGGGGAGGGGCTTGAAGGCGCTTTGTGTGGGTGTCTGACCCCATCAGA 92462
OY 246 -----
Db 92463 GAGGCTGTGATGAGATCTTTCGCCCAACCCCAACCTTACGCTCAGTGCAGCA 92522
OY 246 -----
Db 92523 AGATTACTGCCGAGGCTCTGCTTTTCAGAGTGCAGCCCTAGTACCTCTCTCCCTC 92582
OY 247 -----HisValLeuHisLeuGluProGlySerValGlySerGlyArgCysPro 264
Db 92583 TATGTAGATGTCTCCACTGATCTCAGCAATGCAATGAAATGAAAGGGGCGTGTGCC 92642
OY 264. onIAGLubProSerArgProPheAlaSerThrPhe----- 275

Db 92643 ACACGAGCCGCGCTCCCTTGTGCGCAGTACTTTTGTAGTGGGTGATACCAACCCAG 92702
OY 275 -----
Db 92703 TGGCAGAGCGTAGAGTGTATCTTACTTCCGGTATGGATGGCGCTGTCTTAAGCC 92762
OY 276 -----
Db 92763 TGGCCCTCTCTCTGACTGTGCGCCCTTCTAGGTGGGAGCTGTACAGGGCTCACTCG 92822
OY 286 AspPheLeuGlyArgGlnAlaMetIlePheArgSerGlyProArgProAlaLeuArg 306
Db 92823 TGAATTCCTGAGCAGTGAAGCCATGATTTTCCGAGTGGGGGTCCCGACAGCCCTGCG 92882
OY 306 gSerAspSerAspGlnSerLeuHis-----
Db 92883 TTCTGACTCTGACCAAGAGGCTCTCTACATGTGAGCTCCCTGAGGGGTGAGGTGCTTGG 92942
OY 315 -----
Db 92943 ACCATCTCTTACCTGGAGAGAGAGATAGATGAATGTGGCCAAAGCAGCGGC 93002
OY 315 -----
Db 93003 CACTGTCTCAGGTCTGGGGTGGGGAGCGCTGTGCTCCCTGCATCATTAACCC 93062
OY 316 -----AspProArgPheValMetAlaAlaArgIleProGlyLysSerAspGln 331
Db 93063 TGTGACCTTCTCAAGATCCCGGTTGTGATGGCTGTGCGATCCCAATATCTCCAGC 93122
OY 331 nAspAsnAspLysValTyrPhePhePheSerGluThrValProSerProAspGlyLys 351
Db 93123 GGATGATGACAAAGTGTACTTCTTCTGTGAGCTGTCTTCAACAGATGTGGCC 93182
OY 351 rAsnHisValIleThrValSerArgValGlyVal-----
Db 93183 AGGTATGTACACCGTACCGCGGTGGTGTGAGCTGTGTGTGAGAGAGTGTGGGGTG 93242
OY 364 -----
Db 93243 GTGAGATTCAAGACCCGTGTGTGACTTCTCCCTAGACTCTCCCTTCACAGTAGAGA 93302
OY 364 -----
Db 93303 ACTGAGCAAGGATACAGCGCTACATACATCCCTCTTGTGTGTGAGGTAGAGGA 93362
OY 365 -----
Db 93363 CAGAGGCGAGGGCCATGTCTACAGATCATAAAGTGACTCCACCTCCATCCAGATATG 93422
OY 366 sPalaGlyGlyGlnArgValLeuValAsnLysTyrPheThrPheLeuLysAlaArgLeu 386
Db 93423 ATGCTGTGGCGACGGGTGTGTGATTAATGAGACAGCTTCTCAAGGCCAGCTGG 93482
OY 386 aLysSerValProGlyProGlyValAlaGluThrHisPheAspGlnLeu-----
Db 93483 TATGCTGTGTGCTGGCGCTGCGGGGCTGAGACTACTTGAACACAGCTGGGTAAAGGA 93542
OY 402 -----
Db 93543 TGGCCAGAGGGCTAGATCTGGATGGGAGAGGCTCAGAGCTGTGAATCAGTTTATG 93602
OY 403 -----
Db 93603 GGCAGTGGCTATCTATGATATCTATGCTTTCTGTGTCTCCAGAGAGATGTCTCTGG 93662
OY 409 TrpProLysAlaGlyLysSerLeuGluValTyrAlaLeuPheSerThrVal-----
Db 93663 TGGCCAAAGGCGAGAAACCTCGAGGATGTATGGCTGTTCAGACATGTGAGTGGGCA 93722
OY 425 ----- 425

```


US-09-813-290-2 (1-875) x AF034744 (1-2898)

QY ArgGlnGlyPro-----SerLeuLeuSerSerAlaProLeuProAla 50
 Db 126 AGACGGCGCTCTTGAGACGGCTTCTGCTCCGCCCTTCAACACCGGGCCCAAGACAGA 185
 QY 51 GlnAsp---TrpValGluProLeuProTyrLysTrpTrpProGlyLysSerAlaAsn 69
 Db 186 AAGGCTAGACGGATCCAAATATGTGCCGCAATGACACTTGGGAATGATTTCTGAT 245
 QY 70 TyrAsp-----ArgArgProAlaGlyProGlyLys----- 80
 Db 246 GACAAACCTCTTCTTTGTGACAAACCTGTGCGCCGAGTTGCCCTGAGAGGAATAC 305
 QY 81 -----SerAlaGlyArgArg-GlnArgCysProGlnPheTrp 92
 Db 306 TAAGTAAACTCAATCTCTTAAGTGTGGCTGACGGGGCCAGAGAGACGACACAG 365
 QY 92 OserMetAlaProSerAlaTrpAlaIleCysTrpLeuGlyGlyLeuLeuLeuIleG 112
 Db 366 CACCATGGCAGCGCGGACACATCTCTGCTGCTGGGTGCTGCTGCTGCTGCTGCTG 425
 QY 112 YGlySerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTrpArg 132
 Db 426 CTGAGACCCGAGTCACTCCGCGACACCTCTCTACCCAGGCTACGCTCTCAATPAA 485
 QY 132 PLeuLeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuLeuG 152
 Db 486 ACTTTGGAACTGAATGAGCATTCATATTTCAAAGCCCCCTTGGATTTCTTGATGTC 545
 QY 152 PAlaMetTrpLeuAspGluTrpArgAspArgLeuPheLeuGlyGlyLeuAspAlaLeu 172
 Db 546 TACATTCGTCTGCTGATGATATCAAGACGGCTTTGTGGAGAGGAGACCTTGCTCA 605
 QY 172 rSerLeuArgLeuAspGlnAlaTrpProAspProArgGluValLeuLeuProGlnPr 192
 Db 606 TTCCCTGAACCTTGAAGACAGTCACTGACGGCTACAGAGATATCTGCGCCAGACAG 665
 QY 192 oGlyValArgGluGlyCysValArgGlyArgAspProLeuTrpGluCysAlaAsn 212
 Db 666 AGTAAAGGTAGAAAGATCATATGAAAGAAAGAAC---GCCAAATGGTGTGCCAAATTA 722
 QY 212 eValArgValLeuGlnProHisAsnArgTrpHisLeuLeuAlaCysGlyTrpGlyAla 232
 Db 723 TATCCGGGTTTGGATCACTCAACAGACACACCTTCTGACCTGTGCTGCTGCTGCT 782
 QY 232 eGlnProThrCysAlaLeuIleThrValGlyHisArgGlyGluHisValLeu---His 251
 Db 783 TGATCCACACTGTGCTTCATCAGAGTGGGACCATTCAGAGGAACCCCTGTTTCACT 842
 QY 251 uGlnProGlySerValGlnSerGlyArgGlyArgCysProHisGluProArgProPh 271
 Db 843 GGAGTCACACAGATCTGAGAGAGAGAGGAGAGATGCTCTTTGACCCCACTCTCTCT 902
 QY 271 eAlaSerThrPheIleAspGlyGluLeuTrpTrpGlyLeuThrAlaAspPheLeuGly 291
 Db 903 TGTGTCCACGCTAGTGGGAATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
 QY 291 gGluAlaMetIlePheArgSerGlyArgProArgProAlaLeuArgSerAsp---Ser 310
 Db 963 AGACTCCGCGATCTTCCGCGACATGGGAACTTGAAGCCATTTGCGACGTGAGCATG 1022
 QY 310 pGlnSerLeuLeuHisAspProArgPheValMetAlaAlaArgIleProGluLeuSer 330
 Db 1023 TGAGCGGCTCTCTGAAGAACCAAAATTTGTAGTTCAATATGATTTCTGATTAACGA 1082
 QY 330 pGlnAspAsnArgValArgTrpPhePhePheSerGluThrValProSerProAspGly 350
 Db 1083 CCGAGATGACAAATAATGTAATTTCTTTACTAGAGAAGCGCTGAGGCGGAGAACAA 1142
 QY 350 ySerAsnHisValThrValSerArgValGlyArgValCysValAsnArgIleGly 370
 Db 1143 CGGC---CACACATCTACACCCGAGTGGGCGGCTGTGCTGCTGATGATGAGAGACA 1199

QY 370 nArgValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValPr 390
 Db 1200 GAGAAATCCCTGGTAACAAAGGAGACACTTCTTAAGCGCGGCTGGTTGCTCAGTGC 1259
 QY 390 oGlyProGlyValAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeuTrp 410
 Db 1260 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
 QY 410 oLysAlaGlyLysSerLeuValValTrpAlaLeuPheSerThrValSerAlaValPhe 430
 Db 1320 CAGAGATCTCTTAAGAAATCCAGATGATATTGAGCTGTTAATACACAGCATATTTAG 1379
 QY 430 nGlyPheAlaValCysValTrpHisMetAlaAspIleTrpGluValPheAsnGlyPro 450
 Db 1380 AGGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1439
 QY 450 eAlaHisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyGlyLysValProPhe 470
 Db 1440 TGTCTAATAAGAAAGCGCTGAATACACCTGCTCATATGAAAGAAAGTCCCTTACC 1499
 QY 470 oArgProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySer 490
 Db 1500 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550
 QY 490 rLysAspTrpProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrp 510
 Db 1551 CAAGATTCACCCGATGACGCCATCCGTTCCGAGAGATGATCTCTAATGATATCAG 1610
 QY 510 oValArgProArgHisGlyArgProValLeuValLysTrpHisLeuAlaGlnLeu 530
 Db 1611 CATAAAGCTGTTCAATATAAAGCAATACTGTAATAAAGCAATGGAATACACCTGAG 1670
 QY 530 sGlnIleValValAspArgValGluAlaGluAspGlyTrpTrpAspValIlePheLeu 550
 Db 1671 GCACTTCCCGGTGATCGGCTGAGAGCGAGAGTGGCCAGTGTGACGCTTATTATTTG 1730
 QY 550 yThrAspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlyGlySerAlaGlu 570
 Db 1731 GACAGACAGAGAAATGCTGCAAGATATCATATTTCACCAACCAAGAAACAGATG 1790
 QY 570 oGluGluValValLeuGluGluLeuGlnValPheLysValProThrProIleThrGlu 590
 Db 1791 GGAGAGATGATCTTAGAGAACTTCAATATVCAAGATCCAGGCCCTATATTCTTAT 1850
 QY 590 tGluIleSerValLysArgGlnMetLeuTrpValGlySerArgLeuGlyValAlaGlu 610
 Db 1851 GGAATTTCTTCAAGAGACAAACACCTTTACATTGATCAGCTCTGCTGCTGCTGCT 1910
 QY 610 uArgLeuHisGlnCysGluTrpTrpGlyThrAlaCysAlaGluCysCysLeuAlaArg 630
 Db 1911 CAGATTCATCACTGCGACATGATGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTG 1970
 QY 630 pProTyrCysAlaTrpAspGlyAlaSerCysThrHisTrpArgPro-----SerLeu 648
 Db 1971 CCCGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2030
 QY 648 yLysArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeu 668
 Db 2031 AAAGAGGAGGTTCCGACAGGAGAGCTTCCGATGCGCAAGCGGCCCAAGGTGTTT 2090
 QY 668 yGlnSerGlnGluGluValValGlyLeuValAlaAlaThrMetValTrpGlyThr 688
 Db 2091 ACAGCAATTTTGTGAGCGCGGTTGGACAGACTGAAAGAGGCTGTGCTTATGCAATGA 2150
 QY 688 uHisAsnSerThrPheLeuGluCysLeuProLysSerProGlnAlaValAlaArgTrp 708
 Db 2151 GAGCAACAGTACTCTGTTGGAATGACACCCGCGATCACTAACAACAAAGTCACTGCT 2210
 QY 708 uLeuGlnArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeu 728
 Db 2211 TGTACAGAGAGGACCGCATTAAGAAAGAGAGGTGAAGAGATGACAGATTTGTCAA 2270

```

IRES      Location/Qualifiers
           1..2328
           /organism="Mus musculus"
           /strain="B6B1/c FC3H"
           /db_xref="taxon:10090"
           /cell_line="66C1A"
CDS
           1..2328
           /codon_start=1
           /product="semaphorin H"

```

Qy	94	MetaIaProSerAlaIatrpAlaIleCysTrpLeuLeuGlyIleuLeuLeuHisIsgIyI	113
		::: ::	
Db	1	ATGGACACGGCGGACACATCTCCACCTTCCTCTCGTGGGCTACCTCGTGAAGACTCGG	60
Qy	114	SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerGlyArgAspLeu	133
		::: ::	
Db	61	ACCCAGGCTACCTCCGGACACCCCTCTACCCCGAGGTCAGCCCTGTCACATAAAGACTT	120
Qy	134	LeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAla	153
		::: ::	
Db	121	TTCGAACATGAATAGACTTCATTAATTTCAAGCCCTTCGATTTCTTGATCTCCATACA	180
Qy	154	MetIYrLeuAspGluTrpArgAspArgPheLeuGlyIleuAspAlaLeuYrSer	173
		::: ::	
Db	181	ATGCTGCTGGAGATATCAAGAAAGGGCTCTTTGTGGAGGACAGAGACCTTGCTATTCG	240
Qy	174	LeuArgLeuAspGlnAlaIatrpProAspProArgGluValLeuTrpProProlnProGly	193
		::: ::	
Db	241	CTGAACTTGGACGAGTCGACGTCAGCGCTACAGAGAGATATCTGGCCACACAGCAGTA	300
Qy	194	GlnArgGluGluCysValArgIsgIyArgAspProLeuTrpGluCysAlaAspPheVal	213
		::: :: :: :: ::	
Db	301	AAGGTGAAGAAATGCATATATGAAGAAAGAC--GCAATAGAGTGTGCCATTATATC	357
Qy	214	ArgValLeuGluProHisAsnArgTrpHisLeuLeuAlaCysGlyTrpGlyAlaPheGln	233
		::: ::	
Db	358	CGGGTTTTCATCAGACACAAAGAGACACACCTTGACCTTGCTAGCGAGGCTTTTGAT	417
Qy	234	ProTrpCysAlaLeuIleTrpValGlnHisArgGlyGlnHisValLeu--HisLeuGln	252
		::: ::	
Db	418	CCACACTGTGCTTCCTTCAGAGTGGGGCCACCTTCAGAGAGAACCCCTTGTACCTGGAG	477
Qy	253	ProGlySerValIleuSerGlyArgGlyArgCysProHisGlnProSerArgProPheAla	272
		::: ::	
Db	478	TCACACAGATCTGAGAGAGAGAGGGCCATGTCCTTTTGAGACCCCAACCTCCTCTCTGTG	537
Qy	273	SerTrpPheIleAspGluLeuLeuYrTrpGlyLeuTrpHisAlaAspPheLeuGlyArgGln	292
		::: ::	
Db	538	TCCACGCTACTTGGAGATGAGCTGTGTGTGAGACTCTACAGTACATATTGGGCGCAGAGAC	597
Qy	293	AlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAsp--SerAspGln	311
		::: ::	
Db	598	TGGCGAGATCTTCGACAGCATGGGAGAGTTAGGCGCAATATTCGCACTGACAGATGACGATGAG	657

Oy	312	SerLeuLeuHisAspProAlaGlyPheValIleMetAlaAlaArgIleProGluAsnSerAspGln	331
Db	658	CGGCTCTGGAAAGAACCAAAATTTGTGAAGTTCATATGATTCCATGCCAATACGAAGACCGA	717
Oy	332	AspAsnAspLysValTyrrPhephePheSerGluThrValProSerProAspGlyGlySer	351
Db	718	GATGACACAACAAATGTACTTTTTCTTTACTAGAAGGGCGGTGGAGCGGCAGAACAAACGCC	777
Oy	352	AsnHisValThrValSerTrpArgValGlyArgValCysValAsnAspAlaGlyGlyGlnArg	371
Db	778	---CACACGATCTACACC CGAGGTGGGGGGCGTGTCGTGAATGACATCGCGGACACAGA	834
Oy	372	ValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGly	391
Db	835	ATCCTGGTGAACAAGTGGAGCATCTTCCCTTAAGCGCGGTGGTTCTCAGTCCCGGA	894
Oy	392	ProGlyValGlyAlaGluThrHisPheAspGlnLeuGluAspValIleLeuLeuTrpProLys	411
Db	895	ATGAATGGATCGACACATCTTATGACAGACATAGAGGTGTGTTTTATACGCCGACAGA	954
Oy	412	AlaGlyLysSerLeuGluValTyrrAlaLeuPheSerThrValSerAlaValPheGlnGly	431
Db	955	GATCCTTAAGAAATCCAGTCGTATTTGGCTGTTTAATACACCATATATTATTAAGGC	101
Oy	432	PheAlaValCysValTyrrHisMetLysAspLerTrpGluValPheAsnGlyProPheAla	451
Db	1015	CATGCGTATGTGTATACATGTCATGTCACATGTACCGGAAGCGTTTAATGGCCATATGCT	107
Oy	452	HisArgAspGlyProGlnHisGlnTrpLysProTyrGlyGlyLysValProPheProArg	471
Db	1075	CATTAAGAAGGCCCTGGAATTACACCGTGTACATATATGAAGAAAAGCCTTCACCAAG	113
Oy	472	ProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLys	491
Db	1135	CCGTGTCCTGTCGCCAGAAAGTAAAC-----GGAGGCAAGTAGTAAGAACCCAAAA	118
Oy	492	AspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProVal	511
Db	1186	GATTACCCCCGATGAGCCATCCGGTGTGGCAAGATCGCATCTCTAATGTATAGACCCATA	124
Oy	512	ArgProArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnIleuHisGln	531
Db	1246	AACCTGTTCAATFAAAAAACCAATCTGTAAAAACAGATGGAAAAATCAACCTGGAGCAA	130
Oy	532	IleValValAspArgValGluValAlaGluAspGlyThrTyrrAspValIlePheLeuGlyThr	551
Db	1306	CTTGCCCGTGATCGGGTGGAGCGGAGAGATGGCCAGATAGACCTTATTATTATGGGACA	136
Oy	552	AspSerGlySerValLeuLysValIleAlaLeuAlaGlyLysSerAlaLeuProGlu	571
Db	1366	GACACAGGAATTTGTGTGAAGATATCACATTTTACACCAACCAAGAACAGATGGATGGAG	142
Oy	572	GluValValLeuGluGluGluLeuGlnValPheLysValProThrProIleThrGluMetGlu	591
Db	1426	GAACTCATTTAGAGGAATCTCAAAATTTTTCACAGATCAGCCCATCATCTTTCTATGGAA	148
Oy	592	IleSerValLysArgGlnMetLeuTyrrValGlySerArgLeuGlyValAlaGlnLeuArg	611
Db	1486	ATTCTTCCAAGAGACAAACAGCTTTTACATTTGGATCGATCGCTCGCTGGCACMACTGAGA	154
Oy	612	LeuHisGlnCysGluThrTyrrGlyThrAlaCysAlaGluCysCysLeuAlaArgAspPro	631
Db	1546	TTCATCTACTGCAACATGTATGGCAATGCTTGTGTGATGACTGCTGCTGCTGCGAGACCG	160
Oy	632	TyrCysAlaTrpAspGlyLaseSerCysThrHisTyrrArgPro-----SerLeuGlyLys	649
Db	1606	TACTGTCCCTCGGATGGCATATCTCTGCTCCAGTACTACCCAAAGGTGCACACGCCAAG	166
Oy	650	ArgArgPheArgArgGlnAspLierArgHisGlyAsnProAlaLeuGlnCysLeuGlyGln	669
Db	1666	AGGAGGTCCCGCAGGAGACGTGTGGCATTTGGCAACGCCGCCAAACATGTGTTTGGACAG	172

OY	670	SerGlnIngluGIuValIgLLeuValAlaIalThrMetValTyrcylThrcLuhis	689
Db	1726	CAATTGTGGAGACCGCTTGCAGCAGACTGAAGAAGCGCTGGCTTTATGTAAGAC	1785
OY	650	AsnSerThrPheLeuGluCysLeuProLySserProGlAlaAlaValAlarPrleuLeu	709
Db	1786	AACGATCTCTGTGTGAATGTCACCCCGCCGATCTACAAAGCAAAAGTCACTGGTTTGTA	1845
OY	710	GlnArpProGlyAspGluGlyProAspGlnValLysThrAspGluValLeuHisThr	729
Db	1846	CAGAAGGACCGCACGTAAAGAAAAGAGAGGTAAACGCGAAGACAGATTGTCAAMGATG	1905
OY	730	GluArgGlyLeuLeuPheArgGlaSerArgPheAspAlaGlyThrTyThrCysThr	749
Db	1906	GACTTGGGCTTGCTCTTCCTCACAAGTACCAAGTCAAGTCAAGGACCACTATTTTGGCAG	1965
OY	750	ThrLeuGlnHisGlyPheSerGlnThrValAlaArgLeuAlaLeuValIleValAla	769
Db	1366	ACAGTAAACAACAATTGTGTCCATACTCTGGCTAAACACCTTGAGAGTGTCGAAGAG	2025
OY	770	SerGlnLeuAspAsnLeuPheProFrogIuProLySproGluGlu-----	784
Db	2026	CATAAATGGAGGGCATGTTTCATAAGACCATGAAGAGGAAGACATCACAGATGCC	2085
OY	785	---ProProAlaArgGlyGlyLeuAlaSerThrProProLySalatryrTLyAspDile	803
Db	2086	TGCCCTCCCTTAAGCGGTATGTCTCAGGGGACAAACCG-----TGSTNCAAGGATTC	2139
OY	804	LeuGlnLeuIleGlyPheAlaAsnLeuProArgValAspGluTyrcysGluArgValTrp	823
Db	2140	TTGCAGCTGATTGGCTACACGCAACTTCCAGAGAGTGAAGAACTCTCGAANAAGCTGTG	2199
OY	824	CysArgGlyThrThrGluCysSerGlyCysPheArgSerArgSerArgGlyLysGlnAla	843
Db	2200	TGTACAGAT-----AAGAAGAGAGAAAAAGCTTAAAGTCTC	2235
OY	844	ArgGlyLysSerTrpAlaGlyLeuGluLeuGlyLysMctLysSerArgValHisAla	863
Db	2236	CCCTCCAAGTGAAGATGATGCCAACCCCGAGAAAAGAGCGTTCGCTCTAAA-----GCT	2289
OY	864	GluHisAsnArgThrProArg	870
Db	2290	GAGCACTTCGCCCTGCCGAG	2310
RESULT 13			
AF022947			
LOCUS	AF022947	2615 bp	mRNA linear VRT 14-OCT-1997
DEFINITION	Gallus gallus collapsin 5 mRNA, complete cds.		
ACCESSION	AF022947		
VERSION	AF022947.1	GI:2522205	
KEYWORDS	.		
SOURCE	chicken.		
ORGANISM	Gallus gallus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
Phasianinae; Gallus.			
1 (bases 1 to 2615)			
Felner,L., Koppel,A.M., Kobayashi,H. and Raper,J.A.			
Severed Chick Semaphorins Bind Recombinant Neuropilin with Similar			
Affinities but Bind Different Subsets of Neurons In Situ			
Neuron 19 (1997) In Press			
2 (bases 1 to 2615)			
Felner,L., Koppel,A.M., Kobayashi,H. and Raper,J.A.			
Submitted (05-SEP-1997) Neuroscience, Univ. of Pennsylvania, 105			
Johnson Pavilion 36th and Hamilton, Philadelphia, PA 19104, USA			
Location/Qualifiers			
1..2615			
/organism="Gallus gallus"			
/db_xref="taxon:9031"			
22..2379			
/codon_start=1			
/product="collapsin 5"			
FEATURES			
Source			
CDS			

/protein_id="AB80952.1"
 /db_xref="GI:2522206"
 /translation="MLGRMASADLLILALCGILLELPAGYATDTPORPLRSHKL
 MDNRPSVSPRGVYLHMLDEYOEFLVSGRDLVLSADRLSNRRLHMSST
 PLOAECTIKGRADDECAANVRLHRYNRHLLACGMDPACCTEFLRVNRAHLFT
 QLESNHFERGRKCPDPPTSSFTILGELFTGLSDVGRNAAVEPRTRNRAHLFT
 EPDSHLEKPKFVSGYMTIPDNEDHDDNKVYLFPEKALAEETSTHATITRYGRVYCN
 DMGGRIVANKMSTFLKTRLVCSVGRNGIDTFHDELEDFLLQTRKNRPVLEGLFS
 GLTNIFRGYALICYHMAIVRAENPGVNGHGEYVWALYEGVPPRPGSCAKVNG
 QVLTTRKDYDEAVHFARSHPLMOPKIPYHRRPILVKTGKYLKQIAVDREAEDG
 IGTRESDYAOVKFHOCMDYGTACADCCALPCYAMQISGRVPTGMOAKRRRRDD
 VRNGNAQQCEGQOFIGEVLEKTEBERLYVTEINSTLLETPTLQAKVMVQYRAE
 TKREVKTDERRILKMDGLFLKLRHLDAGTYFOVLEHSTVHTVRKITLEIEEYV
 DEMSKDYEEELSHKMPQMSNPQVSKPMYKEFLQIGISNFRVREYCEKVMCTD
 KKKRKLKMSPSKMYKAPQEKRODOEKARIRPEHYRLPRNIADS"

BASE COUNT 802 a 518 c 620 g 675 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 1,93e-73 Length: 2615
 Score: 2033.50 Matches: 401
 Percent Similarity: 65.13% Conservative: 122
 Best Local Similarity: 49.94% Mismatches: 251
 Query Match: 42.85% Indels: 29
 DB: 5 Gaps: 11

US-09-813-290-2 (1-875) x AF022947 (1-2615)

QY 84 ARGATRGALNARCYSPROGLINHEPROSERMETALAPROSERALATPRLALIECYSTRP 103
 DB 4 AAGAGACCAAGATTTCATCTGGGACAGATGCATCGCCACGACCTTTAATTCG 63
 QY 104 LEULEUGLYLEULEUENHISGLYISERSERGLYPROSERPROGLYPROSERVAL 123
 DB 64 GCACATATGCGCTTGTACTAGATTACCGCTGGATATCACGCCAACATACGACAG 123
 QY 124 PROATRGLEUARGLEUSERYRARGASPLEULEUSERALASARGSERALALIEPHELEU 143
 DB 124 CCAGAGTTACGCTGTGCATATAAGAGCTCTGGAGTTAAACGAGACGTCAGCTTCAT 183
 QY 144 GLYPROGLIGLYSERLEUENHISGLYISERSERGLYPROSERPROGLYPROSERVAL 163
 DB 184 AGCCCATTTGGGATCTTGGTCTTCAATATATCTCTCGATGAATATCAAGACGACTG 243
 QY 164 PHELEUGLYLEULEUENHISGLYISERSERGLYPROSERPROGLYPROSERVAL 183
 DB 244 TTCGTGGAGAGAGACACCTCTTGTACTCCCTCGATGATGATGATGATGATGATGAT 303
 QY 184 ARGGLUVALLEUENHISGLYISERSERGLYPROSERPROGLYPROSERVAL 203
 DB 304 CGAGAGATTCATTTGGCTTACACCTCTTCMAAGCAGAAATGCAATTAATAAGAGAGA 363
 QY 204 ASPPROLEUENHISGLYISERSERGLYPROSERPROGLYPROSERVAL 223
 DB 364 GAT---CGTGAATATGCTAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 224 LEULEUACYSGLYTRPARGALAPHEGLINPROHRCYSALAEULIETNVAIGLHIS 243
 DB 421 CTTCTGCTGTGGAGACAGAGCTTTGATTCACATTTGTACGTTTAAAGAGTTGGACAT 480
 QY 244 ARGGLYGLU---HISVALLEUENHISGLYISERSERGLYPROSERVAL 262
 DB 481 CCATCAACAGAGATCTCTGTTCACTGCAATCGCAAAATTTGAGAGAGCGAGCGAGCG 540
 QY 263 CYSPROHISGLYTRPARGALAPHEGLINPROHRCYSALAEULIETNVAIGLHIS 282
 DB 541 TGCCCTTTTGGCCCT 600
 QY 283 GLYLEUTHRALASAPHELEUGLYARGGLUVALMETILEPHEARGSERGLYGLYPROARG 302
 DB 601 GGTCTCTACATGACTACTGGGAAAGAGATCTGCTGCTTTCGACACATGATGATGATG 660

QY 303 PROALAEUARGSERASPSERASP---GLINSEULEUENHISASPPROARGPHEVALMET 321
 DB 661 GCATATCTCCCACTGCAACCTGATATGACACCTGCGTGAAGAAACCGAAGTGTGTTG 720
 QY 322 ALALAEARGILEPROGLUASNSERASPGLINASPASNAPLYVALTYRPHEPHESEER 341
 DB 721 TCATACATGATTCCTGCAACATGAAAGACCATCATGCAATCAAGATATCTCTTTTACT 780
 QY 342 GLUHTHVALPROSERPROASPGLYGLYISERSERHISVALTHRVASERARGVALYARG 361
 DB 781 GAAAAACGATTAGAGCTGAG---ACANSCACTATGCGCATTTACACAGAGTGGAGCG 837
 QY 362 VALCYSVALASNASPALLAGLYGLINARVALLEUVALASNLVSTRPSETRHPHELEU 381
 DB 838 GTGTGTGATATGATATGAGAGAGACAGCGGATCGTTGAAATAATGAGACCTTCTCTC 897
 QY 382 LYSALAEUARGLEUVALCYSSEVALPROGLYPROGLYGLYALAGLUTHHISPHASG 401
 DB 898 AAACACAGACTAGTTGTCGTGCTGGAGAAATGGATATGATACCATTTTGTATGA 957
 QY 402 LEUGLUSPVALPHELEUENHISGLYISERSERLEUGLVALTYRALAEU 421
 DB 958 TTAGAGATGATGTTTGTGCTGCAACTGATTAACAAAATCCAGTATATTGGCTTC 1017
 QY 422 PHESETRHVALSERALAVALEHGLINGLYPHEALAVACYSVALTYRTHHISMETL 441
 DB 1018 TTCAGCACTACACAAATATATTTAGAGATATGCTATATGCTTATACCATATGGCAATC 1077
 QY 442 ILETTRPGLUVALPHESENGLYPROPHAEALHISARGASGLYPROGLINHISGLIN 461
 DB 1078 GTCCGACAGCTTTTCATATGACCATATGCTCATTAAGAGACCAATAATACCTGGCT 1137
 QY 462 PROTYRGLYGLYLYSVALPROPHAEPROARGPROGLYVALCYSPROSERLYSMET 481
 DB 1138 CTGTACAGAGAGGAAATGACTTATCTAGCCCTGTTCAATGTCCACGAAAGGAAT 1194
 QY 482 GLINPROGLYARGPROPHAEGLYISERSERHISASPLYPROASPGLUVALLEUGLIN 501
 DB 1195 -----GGCGGTGTGATACCTACCAACAGACTACCTATATAGACTGTCATTTTGGC 1248
 QY 502 ARGALHISPROLEUENHISGLYISERSERGLYPROARGHISGLYARGPROVALLEU 521
 DB 1249 AGAGATCATCTAATATGATATGATGACCCATGACCCCTGTTCAATTAAGACCAAT 1308
 QY 522 LYSETHISLEUVALAGLINGLULEUENHISGLINLEVALVALASPARVALGLUVAL 541
 DB 1309 AAAACAGATGGAAAGTACACACTTAAGCAATATGCTGTGACACAGAGTGAAGCTGA 1368
 QY 542 GLYTHRYRSPVALLEPHELEUGLYTHRASPSERGLYSERVALLEUVALLEU 561
 DB 1369 GGGCAATATATGTTGTTGATATGATGACCAATATGATGATGATGATGATGATGAT 1428
 QY 562 LEUGLINALAGLYSERALAGLUPROGLUVALLEUGLUGLUGLUGLUGLUGLUGL 581
 DB 1429 ATTACATATCAAGACAGACAGATCAATGAGAAAGTATTTCTTGAAGAGCTGCAAGT 1488
 QY 582 LYSVALPROTHRPROILETHHGLUMEGLUHLESERVALYSARGGLIMETLEUTRY 601
 DB 1489 AAGGTGCCAATCTCTATCTTCCATGGAATCTCTTCAAAAAGGCAACACCTTACT 1548
 QY 602 GLYSERARGLEUGLYVALALAGLULEUARGLEUENHISGLINCYSGLUHTHRYTH 621
 DB 1549 GGAATGAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
 QY 622 CYSALAGLUCYCYSEULEUVALARGASPPROTYRYSALATPASPAGLYVALSE 641
 DB 1609 TGTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668
 QY 642 HISTYRARGPROSER-----LEUGLYLSARGARGPHEARGARGGLINASP 659
 DB 1669 CGGTACTACCCACAGCAATGACAGCAAAAGAGAGCTTCGAGACAGATGATGATGAT 1728
 QY 660 GLYASPROVALLEUGLINCYSLEUGLYINSERGINGLUGLUGLUGLUGLUGLUGL 679

QY 412 AlaGlySerLeuGluValThrAlaLeuPheSerThrValSerAlaValPheGlnG 431
 DB 1564 GATCTTAAGAAATCCAGTGTATTTGGACTGTTTAACTACCGAACAATATTTAGAGCC 1623.
 QY 432 PheAlaValCysValThrHisMetAlaAspIleTrpGluValPheAsnGlyProPheAla 451
 DB 1624 CATGTGTATGTGTATTCACATGTCAAGTATCCGGAAAGCCTTAAATGCGCCATATGCT 1683
 QY 452 HisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyGlyLysValProPheProArg 471
 DB 1684 CATAAAGAAGGCCCTGAATACACACGTCTATGAAAGAAAGTCCCTACCAACAG 1743
 QY 472 ProGlyValCysProSerIleMetThrAlaGlnProGlyArgProPheGlySerThrLys 491
 DB 1744 CCTCGTTCCTGCGCAGCAAAAGTAAAC-----GGAGGCAAGATGGAACCAACCA 1794
 QY 492 AspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetThrProVal 511
 DB 1795 AGATTACCCGATGACCGCATCCGTTCCAGACGATGCATCTTATATGATCAGCCATA 1854
 QY 512 ArgProAlaGlnHisGlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGln 531
 DB 1855 AACCTGTTCATAAACCAACCAATCTGTAAACACAGATGAAATATCACTGAGGCA 1914
 QY 532 IleValValAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThr 551
 DB 1915 CTTCGCCGTGATCGGGTGAACCGAGATGCGCAGTATGAGCTGTTATTTATGTCACA 1974
 QY 552 AspSerGlySerVal--LeuLysValIleAlaLeuGlnAlaGlyLysValAlaGluPro 570
 DB 1975 GACACAGCAATTTGCTGCTGAAGTAATACAAATTTACACCAAGAAACAGATGGATG 2034
 QY 571 GluGluValValLeuGlnGluLeuGlnValPheLysValProThrProIleThrGluMet 590
 DB 2035 GAGGAAGTCAATCTAGAGAACTTCAAAATATTCAGATCCAGCCCTATCATTTCTATG 2094
 QY 591 GluIleSerValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeu 630
 DB 2095 GAAATTTCTTCAAGAGACAAACAGCTTTACATTTGATCAAGCCCTGCTGTGGCAAGTG 2154
 QY 611 ArgLeuHisGlnCysGluThrTyrGlyThrAlaCysAlaGlnCysCysLeuAlaArgAsp 630
 DB 2155 AGATTCCATCATCTGCGCATGTATGCGCATGCTGTGCTGACGTGCTGCTGCTGCGAC 2214
 QY 631 ProTyrCysAlaThrAspGlyAlaSerCysThrHisTyrArgProSerLeuGly----- 648
 DB 2215 CCGTACTGTGCTGGATGGCATATCTCTGCTCCAGTACTACCCAAAGGTGCACAGCA 2274
 QY 649 LysArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGly 668
 DB 2275 AAGAGAGAGTCCCGCAGGAGGACGTTGGCATGCGCAACGCCCAACAGTCTTTGG 2334
 QY 669 GlnSerGlnGluGluValAlaGlyLeuValAlaIleThrMetValTyrGlyThrGlu 688
 DB 2335 CAGCAATTTGTTGAGACCCGTTGGCAGAGCTGAAGAGAGCGCTGTATGCGCATAGAG 2394
 QY 689 HisAsnSerThrPheLeuGlnCysLeuProLysSerProGlnAlaIleAlaValArgTrpLeu 708
 DB 2395 AGCAACAGTACTGTGTGGAATGCACCCGCTATCACTACAAAGAAAGTCAATCTGTTT 2454
 QY 709 LeuGlnArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeuHis 728
 DB 2455 CTACAGAAAGGAGCGAGCGAAGAAAGAAAGAGGTGAAGCGATGACAGAGTTGTCAAG 2514
 QY 729 ThrGluArgGlyLeuLeuPheArgGluArgLeuSerArgPheAspAlaGlyThrTyrThrCys 748
 DB 2515 ATGGACTTGGGCTTGTCTTCCACAGATACCGAAGTCAAGATCAGGACCTATTTTTC 2574
 QY 749 ThrThrLeuGlnHisGlyPheSerGlnThrValArgLeuAlaLeuValIleVal 768
 DB 2575 CAGACAGTACAAACAATTTTGTCCATCTGTGCTGAATAATCACTTGGAGGTGTCGAA 2634

QY 769 AlaSerGlnLeuAspAsnLeuPheProProGluProLysProGluGlu----- 784
 DB 2635 GAGCATTAAGTGTGAGGCGCATGTTTCAATAGACCAATGAAGAGAAAGACATCAAGATG 2694
 QY 785 -----ProProAlaArgGlyGlyLeuAlaSerThrProProLysAlaIlePtyrLysAsp 802
 DB 2695 CCGTGCCCTCCCTTAAGCCGCTATGCTCAGGGGACAAACCG-----TGGTACAGAGAA 2748
 QY 803 IleLeuGlnLeuIleGlyPheAlaAsn--LeuProArgValAspGluThrCysGluArg 821
 DB 2749 TTCTTGCACTATTTGGTACACAGCAGCAGAGTTCAGAGAGTGAAGATATCTCGAAGA 2808
 QY 822 ValTrpCysArgGlyThrThrGluCysSerGlyCysPheArgSerArgSerArgGlyLys 841
 DB 2809 GTGTGTTATACAGAT-----AAGAAAGCAAAAGCTTAA 2844
 QY 842 GlnAlaArgGlyLysSerTrpAlaGlyLeuGlnLeuGlyLysLysMetLysSerArgVal 861
 DB 2845 ATGTCTCCCTCCAAAGTGAAGATGCAACCCCAAGAAAGAGGCTGCTTAA 2901
 QY 862 HisAlaGlnHisAsnArgThrProArg 870
 DB 2902 ----GCTGACACTTCGCGCTGCCAGG 2925

Search completed: October 9, 2002, 21:11:08
 Job time : 3203 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 9, 2002, 18:51:59 ; Search time 242 Seconds
(without alignments)
6207.852 Million cell updates/sec

Title:	US-09-813-290-2
Perfect score:	4746
Sequence:	1 MACALAGKVPFPGSWPVNHH.....KMKSRVHAENRTPREVEAT 875

Scoring table:	BLOSUM62	
Xgapop	10.0	Xgapext 0.5
Xgapop	10.0	Xgapext 0.5
Fgapop	6.0	Fgapext 7.0
delop	6.0	Delext 7.0

```
Searched:      1736436 segs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Maximize Match 100%
Listing first 45 summaries

```

Command line arguments:
-MODEL=lfame-p2n.model
-o=/cgn2.1/USPTO_SPOOL/US09813229/runat_09102002_094645_19112/app_query.fasta.1.1033
-BD=N.Geneset_032802 -OFMT=fastap -SUFFIX=ext -MINMATCH=0.1 -LOOPCTL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=humano.0.cbf
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=0509813229.ecn.1.1.74.grunat.09102002_094645_19112 -NCP=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAOP=6 -XGAOP=6 -XGAOP=6 -XGAOP=6 -XGAOP=6 -XGAOP=6
-XGAOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	^g Query Match	Length DB	ID	Description
1	4746	100.0	2628	AAH47790	Novel human protel
2	4746	100.0	3568	AAH47792	Novel human protei
3	4357	91.8	4859	AAA93109	Human secreted pro
4	4215	88.8	2349	AAH47791	Novel human protei
5	4046.5	85.3	2340	AAP00250	Nucleotide sequenc
6	3102.5	65.4	2337	AAF00521	Nucleotide sequenc
7	2056	43.3	2898	AAV3567	Human semaphorin E
8	2056	43.3	4460	AAZ28469	Mouse semaphorin H
9	1955	41.2	3988	AAZ28470	Mouse semaphorin H
10	1834	38.6	2709	AAH7049	Semaphorin D cDNA
11	1830	38.6	2530	AAH75767	Human semaphorin I
12	1830	38.6	2601	AAO67442	Human semaphorin I
13	1721	36.3	3331	AAA86112	Human brain tissue
14	1721	36.3	2871	AAA37109	Human DNA encoding
15	1721	36.3	3871	AAAS4098	DNA encoding prote
16	1721	36.3	3871	AAAF5421	Human brain tissue
17	1721	36.3	3880	AAAG89113	Clone BR533-4 codi
18	1558	32.4	2975	AAAC03792	Human collapsin CD
19	1538	32.4	2275	AAAC6680	Human immune/hema
20	1295	27.3	1481	AAQ93231	Human secreted pro
21	988	20.8	3203	AAK81569	Human secreted pro
22	988	20.7	3199	AAAC61570	Mouse CD100 antige
23	974	20.5	1444	AAAC69627	Mouse CD100 nuclea
24	892	18.8	4391	AAAT60666	Murine CD100 DNA s
25	880	18.8	2769	AAAT7462	Mouse extracellular
26	890	18.8	3293	AAD08048	Human semaphorin e
27	887	18.7	2769	AAV31121	Human semaphorin G
28	887	18.4	4157	AAAT60665	Human semaphorin G
29	874	18.4	4157	AAAT7463	Human semaphorin G
30	874	18.4	4157	AAAT7463	Human semaphorin G
31	874	18.4	4157	AAAT7463	Human semaphorin G
32	873	18.4	3503	AAAC92864	Murine M-sema-F CD
33	869.5	18.3	3781	AAAC37092	Human DNA encoding
34	869.5	18.3	3781	AAAC37092	Human Prolactin (UNO
35	869.5	18.3	3781	AAAC37092	Human DNA encoding
36	867.5	18.3	3781	AAAC37092	Human DNA encoding
37	867.5	18.3	3781	AAAC37092	Human DNA encoding
38	864.5	18.2	2951	ABAO4075	Human semaphorin G
39	861	18.1	2517	ABAO4076	Human semaphorin G
40	861	18.1	2598	ABAO4078	Human semaphorin G
41	857	18.1	2155	AAAC4887	Human SEC6 nucleic
42	857	18.1	2155	AAAC4887	Human SEC7 nucleic
43	857	18.1	2284	AAAC8488	Human SEC7 nucleic
44	857	18.1	2703	AAAH7437	Sequence encoding
45	855	18.0	3556	AAAD08283	Human secreted pro

ALIGNMENTS

RESULT 1
AAH47790 ID
AAH47790 standard; CDNA; 2628 BP.

AAH47790;
07-JAN-2002 (first entry)

Novel human protein (NHP) encoding cDNA sequence.

NHP: novel human protein; secreted protein; semaphorin; oxytocin;
neurohypophysial; nootropic; gene therapy; drug screening; ss.
Homo sapiens.

key Location/Qualifiers
1..2628 /tag- a

/product= "NHP"

W0200170806-A2.

27-SEP-2001.

20-MAR-2001; 2001WO-US08834.

20-MAR-2000; 2000US-190638P.

22-MAR-2000; 2000US-191188P.

31-MAR-2000; 2000US-193639P.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Milgowski NL, Turner CA, Hilbun E, Wang X, Donoho G,
Scoville J.

WPI, 2001-611483/70.

P-PSDB; AAG65619.

New polynucleotides encoding human proteins that share structural similarity with semaphorin proteins, protein hormones of neurohypophyseal family for drug screening, diagnosis and therapy of biological disorders

Claim 1, Page 34-35; 43pp; English.

The invention relates to novel human secreted proteins (NHP) that share structural similarity with semaphorin proteins, protein/peptide hormones of the neurohypophyseal family and oxytocin (neurophysin 1 precursor) family. The NHP nucleotide sequences are useful in drug screening of techniques for treating symptomatic or phenotypic manifestations of perturbing the normal function of NHP in the body. Nucleotide constructs encoding NHP products are useful in gene therapy for modulating NHP expression. The constructs can be used to genetically engineer host cells to express NHP products in vivo, these genetically engineered cells function as bioreactors in the body delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences also find use in molecular mutagenesis/evolution of proteins that are partially encoded by the NHP sequences. The encoded NHP polypeptides are useful for generating antibodies, as reagents in diagnostic assays, for identifying other cellular gene products related to NHP and as reagents in assays for screening for compounds that are useful in the treatment of mental, biological or medical disorders and diseases. The present sequence represents a cDNA encoding a NHP.

Sequence 2628 BP; 459 A; 836 C; 848 G; 484 T; 1 other;

Alignment Scores:

Pred. No.:	7 19e-252	Length:	2628
Score:	4746.00	Matches:	875
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-813-290-2 (1-875) x AHA47790 (1-2628)

QY 1 MetAlaCysAlaLeuAlaGlyValPheProMetGlySerTrpProValTrpHisLys²⁰Db 1 ATGGCTGTGCTCCCTACTGGGAGAGCTTCCCAATGGGAGCTGGCCAGCTGGCACA⁶⁰QY 21 SerLeuHisTrpAlaAsnLysValGluGluAlaAlaGlyLysArgGlnGlyProSer⁴⁰Db 61 AGCTGCACTGGGCGCAACAAGGTGGAGAGAGGAGGAGGTGAGCGCAAGGCCAGC¹⁸⁰QY 41 LeuLeuLeuSerSerAlaProLeuProAlaGlnAspTrpValGluProLeuProTyrLys⁶⁰Db 121 CTCTCTCTCTCTCCGCGCTCTCTCCGCGAGAGCTGGGTGAGGCACTGTTTAA¹⁸⁰QY 61 TrpTrpProGluGlySerTrpAlaAsnLysTyrAsnArgArgProAlaGlyProGluGly⁸⁰Db * 181 TGGTGGCCGTGGTGGAGGAGCAAACTACAACCGCGCGGCGGAGCCAGAGGCGCT²⁴⁰QY 81 SerAlaGlyArgArgGlnArgCysProGlnPheProSerMetAlaProSerAlaTrpAla¹⁰⁰Db 241 TCTTCAGAGGAGCGGCGAGGAGGCTGCTGAGTCCCGCAGATGGCCCTCGGCTGGCC³⁰⁰QY 101 IleCysTrpLeuLeuGlyGlyLeuLeuHisGlyLysSerSerGlyProSerProGly¹²⁰Db 301 ATTTGGCTGGCTGAGGGGCGCTCTGCTCCATGGGGGTAGCTCTGGCCCCAGCCCGCC³⁶⁰QY 121 ProSerValProArgLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArgSerAla¹⁴⁰Db 361 CCCAGTGTCCCGCGCTGGCGCTCTCTACCGAGACCTCTCTCTGCAACCCCTCTGCC⁴²⁰QY 141 IlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArg¹⁶⁰Db 421 ATCTTCTGGGCGCCAGAGGCTCTGAACTCTCAAGCCATGACTAGTATGATCGCA⁴⁸⁰QY 161 AspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaTrp¹⁸⁰Db 481 GACCGCTCTTCTGTGGGTGGCTGGAGCCCTCTACTCTCTGGCGTGAGCAGCATGG⁵⁴⁰QY 181 ProAspProArgGluValLeuTrpProGlnProGlnProGlnArgGlyValArg²⁰⁰Db 541 CCAGATCCCGGAGGCTCTGTGGCCAGCCAGCCAGGAGAGAGAGTGTGTGCA⁶⁰⁰QY 201 LysGlyArgAspProLeuThrGlyCysAlaAsnPheValArgValLeuGlnProHisAsn²²⁰Db 601 AAGGAGAGATCTTGTGACAGATGGCCAACTTCTGCGGGTCTTACAGCTCAGC⁶⁶⁰QY 221 ArgTrpHisLeuLeuAlaCysGlyTrpGlyAlaPheGlnProThrCysAlaLeuLeuTrp²⁴⁰Db 661 CGAGCCAGCTCTAGCCGTGGCACTGGGGCTTCCAGCCCACTGTGCTTCAACA⁷²⁰QY 241 ValGlyHisArgGlyGluHisValLeuHisLeuGluProGlySerValGluSerGlyArg²⁶⁰Db 721 GTTGGCCACCGGGGAGACATGTGCTCCACCTGGAGCTGGAGTGTGAAAGTGGCCGG⁷⁸⁰QY 261 GlyArgCysProHisGluProSerArgProPheAlaSerThrPheIleAspGlyValLeu²⁸⁰Db 781 GGGCGGTCCCTCAGCAGCAGCCAGCCGCTCTTGGCACACCTTATGAGGGAGGTG⁸⁴⁰QY 281 TyrTrpGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetIlePheArgSerGly³⁰⁰Db 841 TACAGGGTCTACCTGCTGACTCTCTGGGGAGAGGAGCAATGATCTTCCGAATGGAGCT⁹⁰⁰QY 301 ProArgProAlaLeuArgSerAspSerAspGlnSerLeuLeuHisAspProArgPheVal³²⁰Db 901 CCTGGCCAGCTCTGCGTTCGACTGTACCAAGATCTCTTGCACAGACCCCGGCTTGTG⁹⁶⁰QY 321 MetAlaAlaArgIleProGluAsnSerAspGlnAspAsnAspLysValTyrPhePhePhe³⁴⁰Db 961 ATGGCCCGCGGATCCCTGAGACCTGTGACCAGAGCAATGACAAAGTGTACTTCTTCTTC¹⁰²⁰QY 341 SerGluThrValProSerProAspGlyGlySerAsnHisValTrpValSerArgValGly³⁶⁰Db 1021 TCGGAGAGGTCCTCCGCGAGATGCTGCTGCAACCATGTGCTGACCGCGGTGGC¹⁰⁸⁰QY 361 ArgValLysValAsnAspAlaGlyGlyGlnArgValLeuValAsnLysTrpSerThrPhe³⁸⁰Db 1081 CGGTCTGCTGATGATGATGCTGGGGGCGCAGCGGGTGTGTGGAACAATGAGACACTTTC¹¹⁴⁰QY 381 LeuLysAlaArgLeuValCysSerValProGlyProGlyAlaGluTrpHisPheAsp⁴⁰⁰Db 1141 CTCAAGCCAGGCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG¹²⁰⁰QY 401 GlnLeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSerLeuGluValTyrAla⁴²⁰Db 1201 CACCTAGAGAGATGTCTGCTGTGGCCCAAGCGGGAAGGCTCGAGGTGTACGG¹²⁶⁰QY 421 LeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCysValTyrHisMetAla⁴⁴⁰Db 1261 CTGTTAGCAACCTCAGTGGCGGTTCACAGGGCTTGGCGTGTGTGTACCATGGCA¹³²⁰

QY 441 AspIleTgPLeuValPheasnGlyProPheAlaHisArgAspGlyProGlnHisGlnTrp 460
 |||
 Db 1321 GACATCTGGAGAGTTTTCACAGCGGCGCTTGGCCACCGAGATGGGCTCCGACCGACAGTGG 1380
 QY 461 GlyProTyrGlyLysValProPheProArgProGlyValLysProSerLysMetThr 480
 |||
 Db 1381 GGGCCCTATGGGGCAAGGGCTTCCCTCGCCCTGGCGCTGGCCCGCCACAGATGACC 1440
 QY 481 AlaGlnProGlyArgProPheGlySerThrLysAspTyrProAspGlyValLeuGlnPhe 500
 |||
 Db 1441 GCACAGCAGAGAGGGCTTTGGCAGACACAGAGACTACCCAGATGAGGTGCTGACGTTT 1500
 QY 501 AlaArgAlaHisProLeuMetPheTrpProValArgProArgHisGlyArgProValLeu 520
 |||
 Db 1501 GCCCGACCGCCACCGCTCATGTCTGTGGCTGTGGCGCTCGACATGGCGCCGCTGCTCTT 1560
 QY 521 ValLysThrHisLeuAlaGlnGlnLeuHisGlnLeuValValAspArgValGlnAlaGlu 540
 |||
 Db 1561 GTCAAGACCCACCTGGCCACAGCTACACAGATCGTGTGGAGCCGCGTGAGGACAGAG 1620
 QY 541 AspGlyThrTyrAspValIlePheLeuGlyThrAspSerGlySerValLeuLysValIle 560
 |||
 Db 1621 GATGGAGACCTCGATGATCTATTTCTGGGAGCTGACTCAGGGCTCTGCTCAAACTATC 1680
 QY 561 AlaLeuGlnAlaGlyLysSerAlaGluProGlnGluValValLeuGlnLysGlnVal 580
 |||
 Db 1681 GCTCTCCAGAGAGGGGGCTGACGTACCTGAGAAAGTGTCTGAGAGAGCTCCACAGTG 1740
 QY 581 PheLysValProThrProIleThrGlnMetGlnLysSerValLysArgGlnMetLeuTyr 600
 |||
 Db 1741 TTAAAGGTGCCACACCTATCTACCCAAATGAGATCTCTGTAAAAAGCAATGCTATAC 1800
 QY 601 ValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnCysGluThrTyrGlyThr 620
 |||
 Db 1801 GTGGGCTCTGGGCTGGGTGTGGCCGCTGCGGCTGCACCAATGTGAGCTTACGGCAGCT 1860
 QY 621 AlaCysAlaGlnCysCysLeuAlaArgAspProTyrCysAlaThrAspGlyAlaSerCys 640
 |||
 Db 1861 GCCTGTGCAGAGTGTGCTGGCGCCGAGCCCACTGTGCTGGAGTGGTGGCTCTCTGT 1920
 QY 641 ThrHisTyrArgProSerLeuGlyLysArgArgPheArgGlnAspIleArgHisGly 660
 |||
 Db 1921 ACCCATACCGCCCGCCAGCTGTGGCAAGCGCGGTCCGCGGAGAGACATCCGCGACGCG 1980
 QY 661 AspProAlaLeuGlnCysLeuGlyLysSerGlnGlnGlnValValValValValVal 680
 |||
 Db 1981 AACCTGTCCCTGCAGTGTGCTGGGCGCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAG 2040
 QY 681 AlaThrMetValTyrGlyThrGlnHisAsnSerThrPheLeuGlnCysLeuProLysSer 700
 |||
 Db 2041 GCCACCATGTGTACGGAGAGAGACACATATGACCTTCCGTGGAGTGGCTGCCAAGTCT 2100
 QY 701 ProGlnAlaAlaValArgTyrLeuLeuGlnArgProGlyLysGlnGlyProAspGlnVal 720
 |||
 Db 2101 CCCCAAGCTGTGTGGCTGTGCTTTCAGAGGCGAGGGAGTGGGGCTGCACAGAGTGTG 2160
 QY 721 LysThrAspGlyArgValLeuHisThrGlnArgGlyLeuLeuPheArgArgLeuSerArg 740
 |||
 Db 2161 AAGAGCGAG 2220
 QY 741 PheAspAlaGlyThrTyrThrCysThrThrLeuGlnHisGlyPheSerGlnThrValVal 760
 |||
 Db 2221 TTTCATGGGGGAGACCTACACCTGACACCTGTGAGAGCATGTGCTTCCAGAGCTGGAGT 2280
 QY 761 ArgLeuAlaLeuValValIleValAlaSerGlnLeuAspAsnLeuPheProGlnPro 780
 |||
 Db 2281 CCGCTGGGCTGTGTGTATTTGTGGCTTACAGCTGGAACAACCTGTTCCTCCGAGGACA 2340
 QY 781 LysProGlnGluProProAlaArgGlyLeuAlaSerThrProProLysAlaTrpTyr 800
 |||
 Db 2341 AAGCCAG 2400
 QY 801 LysAspIleLeuGlnLeuLeuGlyPheAlaAsnLeuProArgValAspGlyTyrCysGlu 820

Db 2401 AAGACATCTCGACGCTCATTTGCTTGCACACTGCGCCGGGGATGAGTACTGTGAG 2460
 QY 821 ArgValTrpCysArgGlyThrThrGlnCysSerGlyCysPheArgSerArgGly 840
 |||
 Db 2461 CGCTGTGTGGAGAGGGGACACACGAGATGCTCAGGGCTGTCCGAGCCGAGCGGAGGC 2520
 QY 841 LysGlnAlaArgGlyLysSerTrpAlaGlyLeuGlnLeuGlnLysLysMetLysSerArg 860
 |||
 Db 2521 AAGCAGGCGCAGGGGCAAGAGCTGGGAGGCTGAGAGCTAGGCAAGAAAGATGAGAGCCGG 2580
 QY 861 ValHisAlaGlnHisAsnArgThrProArgLysValGlnAlaThr 875
 |||
 Db 2581 GTGCATGCCAGACAAATCGACAGCCCGGGAGGTGAGAGCCACG 2625
 RESULT 2
 ID AAH47792 standard; DNA; 3568 BP.
 AC AAH47792;
 AC AAH47792;
 DT 07-JAN-2002 (first entry)
 DE Novel human protein (NHP) polynucleotide sequence.
 DE NHP: novel human protein; secreted protein; semaphorin; oxytocin;
 KW neurohypophyseal; neotrophic; gene therapy; drug screening; ds.
 OS Homo sapiens.
 PN WO200170806-A2.
 PD 27-SEP-2001.
 PF 20-MAR-2001; 2001WO-US08834.
 PR 20-MAR-2000; 2000US-190638P.
 PR 22-MAR-2000; 2000US-191188P.
 PR 31-MAR-2000; 2000US-193639P.
 PA (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G;
 PI Scoville J;
 DR WPI; 2001-611483/70.
 XX New polynucleotides encoding human proteins that share structural
 PT similarity with semaphorin proteins, protein hormones of
 PT neurohypophyseal family for drug screening, diagnosis and therapy of
 PT biological disorders
 PS Disclosure; Page 40-41; 43pp; English.
 XX The invention relates to novel human secreted proteins (NHP) that share
 CC structural similarity with semaphorin proteins, protein/peptide hormones
 CC of the neurohypophyseal family and oxytocin (neurophysin 1 precursor)
 CC family. The NHP nucleotide sequences are useful in drug screening of
 CC techniques for treating symptomatic or phenotypic manifestations of
 CC perturbing the normal function of NHP in the body. Nucleotide constructs
 CC encoding NHP products are useful in gene therapy for modulating NHP
 CC expression. The constructs can be used to genetically engineer host cells
 CC to express NHP products in vivo, these genetically engineered cells
 CC function as bioreactors in the body delivering a continuous supply of a
 CC NHP. A NHP peptide, or a NHP fusion protein to the body. The sequences
 CC partially encoded by the NHP sequences. The encoded NHP polypeptides are
 CC useful for generating antibodies, as reagents in diagnostic assays, for
 CC identifying other cellular gene products related to NHP and as reagents
 CC in assays for screening for compounds that are useful in the treatment of
 CC mental, biological or medical disorders and diseases. The present
 CC sequence represents a NHP polynucleotide sequence.

Sequence 3568 BP; 633 A; 1099 C; 1156 G; 680 T; 0 other:

Alignment Scores:

9,96e-252 Length: 3568
 Score: 476.00 Matches: 875
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-813-290-2 (1-875) x AAH47792 (1-3568)

QY 1 MetaAlaCysAlaLeuAlaGlyValAlaPheProMetGlySerTrpProValTrpHisLys 20
 Db 555 ATGGCTGTGGCTTACCTGGGAAGTCTTCCCAATGGGGAGCTGGCCACATGTGGCAAAA 614
 QY 21 SerLeuHisTrpAlaAsnLysValGluGluAlaAlaGlyLysArgGlnGlyProSer 40
 Db 615 AGCCTGACCTGGGCAACAAGTGGAGAAGAAAGCGGAGGTGGAGCGCAAGGCCCCAGC 674
 QY 41 LeuLeuLeuSerSerAlaProLeuProAlaGlnAspTrpValGluProLeuProTyrLys 60
 Db 675 CTCCTCTCTCTCTCCGCCCTCTCTCCGCCAGACTGGGTGGAGCCACTGCCTTTAAAG 734
 QY 61 TrpTrpProGlyGlySerArgAlaAsnTyrAsnArgArgProAlaGlyProGluGlyGly 80
 Db 735 TGTGGCGCTGGTGGCAGCAGCAGCAAACTACAAACGGCGGCGCAGCGGAGCAAGAGGGCGGC 794
 QY 81 SerAlaGlyArgArgGlnArgCysProGlnPheProSerMetAlaProSerAlaTrpAla 100
 Db 795 TCTGGAGGAGGCGGAGGCGGCTCCCTCAGTTCCCAAGATGGCCCTCGGCGCTGGGCC 854
 QY 101 IleCysTrpLeuLeuGlyGlyLeuLeuLeuHisGlySerSerGlyProSerProGly 120
 Db 855 ATTGCTGGCTGCTAGGGGGCTCTCTCTCCATGGGGAGTCTGGGCCCGCCAGCCCGGC 914
 QY 121 ProSerValProAlaGlyLeuArgLeuSerTyrArgAspLeuSerAlaAsnArgSerAla 140
 Db 915 CCCAGTGTGGCCCGGCTGGCTCTCCACCGAGACTCTGTGTGCCAACCGCTCTGCC 974
 QY 141 IlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArg 160
 Db 975 ATCTTCTGGGCGCCCGAGGCTCTCTTAACCTCCAGGCGCATGTACCTAGATGAGTACCGGA 1034
 QY 161 AspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaTrp 180
 Db 1035 GACCGCTCTTCTGGGGCTGGAGCGCTCTACTCTCTGGGGCTGGAGCGAGCATGG 1094
 QY 181 ProAspProArgGluValLeuTrpProGlnProGlyGlnArgGluGluCysValArg 200
 Db 1095 CCAAGTCCCGGAGGAGTCTGTGGCACCGCGAGCGAGAGAGAGGAGAGTGTGTGCA 1154
 QY 201 LysGlyArgAspProLeuTrpGluCysAlaAsnPheValArgValLeuGlnProHisAsn 220
 Db 1155 AAGGGAAGATCTTGGACAGAGTGGCCCACTGTGGGGGTGTACACCTCACAAAC 1214
 QY 221 ArgTrpHisLeuLeuAlaCysGlyTrpGlyAlaPheGlnProTrpCysAlaLeuLeuTrp 240
 Db 1215 CGGACCCACCTGTAAGCTGTGGCTGGAGCGCTTCCAGCCACCTGTGGCCCTCATACA 1274
 QY 241 ValGlyHisArgGlyGluHisValLeuHisLeuGluProGlySerValGluSerGlyArg 260
 Db 1275 GTTGGCCACCGTGGAGGATGTCTCCACCTGGAGCTGGAGTGGAGAAATGGCCGG 1334
 QY 261 GlyArgCysProHisGluProSerArgProPheAlaSerTrpPheLeuAspGlyGluLeu 280
 Db 1335 GGGCGGTGGCTTACGAGCGCAGCGCTCTTGGCAGACCTTCAATAGCGGGAGCGT 1394
 QY 281 TyrTrpGlyLeuTrpAlaAspPheLeuGlyArgGluAlaMetIlePheArgSerGlyGly 300
 Db 1395 TACACGGGTCTACGCTGACTTCTGTGGGCGAGAGCGCATGATCTTCCGAAGTGGAGGT 1454
 QY 301 ProAlaGluAlaLeuArgSerAspSerAspGlnSerLeuHisAspProArgPheVal 320

Db 1455 CCGCGCGAGCTGTGGCTTCCGACTCTGACAGAGTCTTGGCACAGACCCCGGTGTG 1514
 QY 321 MetaAlaAlaArgIleProGlnAsnSerAspGlnAspAsnAspLysValTyrPhePhePhe 340
 Db 1515 ATGGCCCCCGGATCCCTGAGAACTGTACACAGACATGACAAAGTGTACTTTCTTC 1574
 QY 341 SerGluTrpValProSerProAspGlyGlySerAsnHisValTrpValSerArgValGly 360
 Db 1575 TCGAGAGCGGTCCCTCGCCGATGGTGGCCGAAACCATGTCACTGTACACCGCGTGGC 1634
 QY 361 ArgValCysValAsnAspAlaGlyGlyGlnArgValLeuValAsnLysTrpSerTrpPhe 380
 Db 1635 CGCTGTGCGTGTATGATGCTGGGGCCACCGGTGTGTGTAACAAATGAGACACTTC 1694
 QY 381 LeuLysAlaArgLeuValCysSerValProGlyProGlyValAlaGluTrpHisPheAsp 400
 Db 1695 CTCAGGCCAGGCTGTGTCTGTGGTGGCCGCGGCTGTGTGTAACAAATGAGACACTTC 1754
 QY 401 GlnLeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSerLeuGluValTyrAla 420
 Db 1755 CAGCTAGAGATGTCTCTGTGTGGCCCAAGCGCGGAGAGCCTCGAGGTGTACGGC 1814
 QY 421 LeuPheSerTrpValSerAlaValPheGlnGlyPheAlaValCysValTyrHisMetAla 440
 Db 1815 CTGTTCAGACCGCTCAGTCCGCTTCCAGGGCTTCCGCTGTGTGTACCAATGGCA 1874
 QY 441 AspIleTrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGlnHisGlnTrp 460
 Db 1875 GACATCTGGAGATGTTTCAACGGGCCCTTGGCCACCGAGATGGCCCTGACACCAAGTGG 1934
 QY 461 GlyProTyrGlyGlyLysValAlaProPheProAlaGlyValCysProSerLysMetTrp 480
 Db 1935 GGGCCCTATGGGGGAGAGTGGCTTCCCTCGCTGGGTGTGGCCCGCCAGCAATGAGCC 1994
 QY 481 AlaGlnProGlyArgProPheGlySerTrpLysAspTyrProAspGluValLeuGlnPhe 500
 Db 1995 GCACAGCAGGAGCGCTTGTGGCAGCAGCAAGAGACTACCCAGATGAGGTGTGCAATTT 2054
 QY 501 AlaArgAlaHisProLeuMetPheTrpProValAlaProArgHisGlyArgProValLeu 520
 Db 2055 GCCCGAGCCCAACCCCTATGTCTGTGGCTGTGGCGCTGCACATGGCCCGCTCTCTT 2114
 QY 521 ValLysTrpHisLeuAlaGlnGlnLeuHisGlnIleValValAspArgValGluAlaGlu 540
 Db 2115 GTCAAGACCCACCTAGCGCCACAGACTACCAAGTCTGTGTGGAGCCGCTGGAGGAGAG 2174
 QY 541 AspGlyTrpTyrAspValIlePheLeuGlyTrpAspSerGlySerValLeuLysValIle 560
 Db 2175 GATGGAGCTACGATGTCAATTTCTGGGAGTCACTACAGAGTCTGTGTCAAAAGTCATC 2234
 QY 561 AlaLeuGlnAlaGlySerAlaGluProGluGluValValLeuGluGluLeuGlnVal 580
 Db 2235 GCTCTCAGGAGGAGGCTCAGCTGAGAACCTGAGGAATGGTGTGTGGAGAGCTCCAGGTG 2294
 QY 581 PheLysValProThrProIleTrpGluMetGluIleSerValLysArgGlnMetLeuTrp 600
 Db 2295 TTTAAGTGGCCACACCTATACCGGAATGAGATCTGTGCAAAAGCAATGCTATAC 2354
 QY 601 ValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisLysGlnGluTrpTyrGlyTrp 620
 Db 2355 GTGGGCTGTGGGTGTGTGGTGTGGCCAGACTGGCTGTGCACCAATGTGAGACTACGAGACT 2414
 QY 621 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAlaSerCys 640
 Db 2415 GCTGTGCAAGTGTGCTGTGGCCGGAGCCCATACTGTGTGCTGGAGTGTGCTCTCTGT 2474
 QY 641 ThrHisTyrArgProSerLeuGlyLysArgArgPheArgArgGlnAspIleArgHisGly 660
 Db 2475 ACCCACTACCGCCCAAGCTTGTGGCAAGCCCGCTTCCGCGGAGGAGCATCCGGCAGCG 2534
 QY 661 AsnProAlaLeuGlnCysLeuGlnSerGlnGlnGluGluAlaValGlyLeuValAla 680

DB 2535 AACCTGCCCTCAGTGCCTGGCCAGAGCCAGAAAGAGGAGTGGACTGTGGCA 2594
 QY 681 AATATMETVALTYTGYLTHRCIUHISANSERTHRPHLEUGLUCYSLEUPROLYSER 760
 DB 2595 GCCACCATGTGTACGGACGAGACATACACTTCTTGAGTGGCTGGCCAAAGTCT 2654
 QY 701 PROGINALAAVALARGTTPHEULEUNGINARGPROGLYASPLUGLYPROASPLINVAL 720
 DB 2655 CCCCAGGCTGCTGGCTGGCTGCTTTCAGAGAGCCAGGGGATGAGGGGCTGACCAAGTGTG 2714
 QY 721 LYSTHRASPLIARGLVALLEUHIHSTRHGLUARGGLYLEUHPHEARGHARGLEUSERARG 740
 DB 2715 AAGACGAGACGACGAGTCTTGACACAGAGCGGGGCTGCTGTTCGACAGGCTTACCGCT 2774
 QY 741 PHEASPLAAGLYTHRTYTHRCYSTHRTHRLUENGLUHIHSPHESEGLNTHRAVAL 760
 DB 2775 TTCGATGGGGGACCTACACCTCCACACTCTGGAGCATGGCTTCCAGACTGTGGTC 2834
 QY 761 ARGLEUALALEUVALVALLLEVALAASERGLINLEUASPLANLEUPHEPROPROGLINPRO 780
 DB 2835 CGCCTGGCTGTGGTGGATTTGGCCTCACAGCTGGACAACTGTTCCTCCGAGCA 2894
 QY 781 LYSPPROGLINUPROPROALARGGLYGLYLEUALASERTHRPROPROLYSALATPTTYR 800
 DB 2895 AACCCAGAGAGGCCCCAGCCGAGGAGGCTTGCTTCCACCCCAAGCCAGCTGTGTAC 2954
 QY 801 LYSASPLILEUNGINLEULIEGLYPHEALASPLANLEUPROARGVALASPLIUTYRYSGLU 820
 DB 2955 AAGGACATCTCTGACGATCATTTGGCTTCCCAACTGCCCCGGGGATGAGTACTGTGAG 3014
 QY 821 ARGVALTRPCYSARGLYTHRTHGLUCYSERGLYCYSPPHEARGSERARGSERARGLY 840
 DB 3015 CGCCTGTGGTGACAGGGGACCCAGAAATGCTCAGGCTGCTTCCGAGACCGGAGCCGGGC 3074
 QY 841 LYSGLINALARGGLYLYSSETRTPALAGLYLEUNGILUENGLYLYSMELLYSERARG 860
 DB 3075 AAGCAGGCCAGGGGACAGAGTGGGAGGCTGGAGTACGCAAGATGATGAGAGCCGG 3134
 QY 861 VALHISALAGLUHISASNAIYTHRPROARGGLUVALGLUALATPTT 875
 DB 3135 GTGCATGCCGAGACAAATCGAGACCCCGGAGGTGGAGCCACG 3179
 DB 3135 GTGCATGCCGAGACAAATCGAGACCCCGGAGGTGGAGCCACG 3179
 RESULT 3
 ID AAA93109 standard; cDNA: 4859 BP.
 XX AAA93109;
 AC AAA93109;
 XX 12-JAN-2001 (first entry)
 DT
 XX Human secreted protein coding sequence SEQ ID NO: 17.
 DE Human secreted protein; cytokine; cell proliferation;
 XX Human; secreted protein; cytokine; cell proliferation; autoimmune disorder;
 KM nutritional supplement; immune modulation; autoimmunity disorder;
 KM haematopoiesis regulation; tissue growth; haemostasis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 79..2427
 FT /tag= a
 FT /product= "secreted protein"
 FT sig_peptide 106..144
 FT /tag= b
 FT mat_peptide 145..2424
 FT /tag= c
 XX
 XX WO200049134-A1.
 XX
 XX 24-AUG-2000.
 XX
 PF 18-FEB-2000; 2000MO-US04340.
 XX

PR 19-FEB-1999; 99US-0120680.
 PR 23-APR-1999; 99US-0298733.
 PR 17-AUG-1999; 99US-0149639.
 PR 23-SEP-1999; 99US-0155686.
 PR 01-OCT-1999; 99US-0157247.
 PR 29-NOV-1999; 99US-0167822.
 PR 29-NOV-1999; 99US-0167823.
 PR 15-FEB-2000; 2000US-0298733.
 PA (ALDH-) ALPHAGE INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 PI WPI: 2000-549267/50.
 DR P-PSDB; AAB23609.
 XX
 PT New secreted proteins and polynucleotides encoding them, which are
 PT derived from Homo sapiens, useful for therapy, diagnosis, and research,
 PT as well as nutritional sources or supplements.
 PS Claim 26; Page 249-250; 309pp; English.
 CC The present sequence is the coding sequence for a human secreted protein.
 CC The sequence was isolated from an adult brain cDNA library. The proteins
 CC and coding sequences of the invention can be used in the isolation of
 CC similar genes and proteins, in the elucidation of their function in vivo,
 CC and to treat a number of conditions. It is possible that they may have
 CC uses as nutritional supplements, as cytokine or cell proliferation
 CC factors, in immune modulation, where they may be used to treat immune and
 CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
 CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
 CC have chemokine or chemotactic activity, haemostatic or thrombolytic
 CC activity, or anti-inflammatory activity.
 XX
 SO Sequence 4859 BP; 1082 A; 1326 C; 1493 G; 958 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.16e-230
 Score: 4357.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 91.80%
 DB: 21
 Gaps: 0
 Length: 4859
 Matches: 807
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0
 US-09-813-290-2 (1-875) x AAA93109 (1-4859)
 QY 69 ASNTYASNAIARGPROALAGLYPROGLINGLYSEALAGLYARGARGINARGYS 88
 DB 4 AACTACAAACCGGGGACGAGGAGGCGGCTTGCAGGACGCGGAGCGGCTC 63
 QY 89 PROGINPHERPSEMETALAPROSERALATPRAHILECYSTRPHEULEUNGILYLEU 108
 DB 64 CCTCAGTTCCTCCAGCAWGGCCCTCGGCTGGGCCATTTGCTGGCTGAGGGGCTC 123
 QY 109 LEULENHISGLYGLYSESERGLYPROSERPROGLYPROSERVALPROARGLEUARGLEU 128
 DB 124 CTGCTCCATGGGGGTAGCTCTGCCCCAGGCCCGGCCCAAGTGTGCCCGCTGCGGCTC 183
 QY 129 SETTYRARGSPLEULEUSERALASNAIYSERALILEPHEULEUNGILYPROGLINLYSER 148
 DB 184 TCCACCGAGACCTCTGTCGCCAACCGCTTGCCTTCTGGGCCCCAGGCTCC 243
 QY 149 LEUASPLINLALMETTYRLEUASPLIUTYRARGSPARGLEUPHEULEUNGILYLEU 168
 DB 244 CTGAACTCCAGGCCATGATACAGAGTACGAGACCGGCTTCTGGGTGGGCTG 303
 QY 169 ASPALALEUTYRISERLEUARGLEUASPLINALATPTPROASPLIUTYR 188
 DB 304 GAGCCCTTACTCTCTGCGGCTGGACGAGCAATGCTCCGAGAGTCTGTG 363
 QY 189 PROPROGLINPROGLINARGGLINGLUCYSVALARGLYSGLYARGASPLIUTYRGLU 208
 DB 364 CCACCGAGCCAGGACAGAGGAGAGTGTGTAAGGAAAGATCTTGTACAGAG 423

QY	209	CysAlaAspPheValArgValLeuGlnProHisAsnArgThrHisLeuLeuAlaCysGly	228
Db	424	TGGCCAACTTCTGGGGGTGTACAGCCCTCACACCGGACCCACTGGTACGCTTGGCC	488
QY	229	ThrGlyAlaPheGlnProThrCysAlaLeuIleThrValGlyHisArgGlyGlnHisVal	248
Db	484	ACTGGGGCCCTTCACGCCCACTGGCTGCTCATACAGTTGGCCACCGTGGGGAGCATGTG	543
QY	249	LeuHisLeuGlnProGlySerValGluSerGlyArgGlyArgCysProHisGlnProSer	268
Db	544	CTCCACCTGGAGCTGGCAGTGTGGAAAGTGGCGGGGGGGTGGCCCTCACAGGCCAGC	603
QY	269	ArgProPheAlaSerThrPheIleAspGlyGluLeuTyrThrGlyLeuThrAlaAspPhe	288
Db	604	CGTCCCTTTGGCAGACACTTCTACAGCGGGAGCTGTACACGGGTCTCACTGTGACTTC	663
QY	289	LeuGlyArgGluAlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAsp	308
Db	664	CTGGGGCCGAGAGCCCATGATCTTCGAACTGGAGTCCCTCGGCCAGCTGTGGCTTCGAC	723
QY	309	SerAspGlnSerLeuLeuHisAspProArgPheValMetAlaAlaArgIleProGluAsn	328
Db	724	TCTGACCAAGCTCTGTGCACACCCCGGTTGTGATGGCCGCGAGTCCCTGAGAC	783
QY	329	SerAspGlnAspAsnAspLysValTyrPhePhePheSerGluThrValProSerProAsp	348
Db	784	TCTGACCAAGCAATGACAGGTGACTTCTTCTGGAAGCGGTCCCTCGGCCGAT	843
QY	349	GlyGlySerAsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGly	368
Db	844	GGTGGCTTGAACCATGTACGTACGCGCGGGCCGCTGTGCTGGTGAATGATCTGGG	903
QY	369	GlyGlnArgValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSer	388
Db	904	GGCCAGCGGGTGTGTGAACAANAAGAGCACTTCTCTCAAGGCGAGGCTGGTGTCTCG	963
QY	389	ValProGlyProGlyGlyAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeu	408
Db	964	GTCGCGCGCCTGGTGTGGCCGACACCCACTTTCAGCCAGTAAGAGATGTTCCTGCTG	1022
QY	409	TyrProLysAlaGlyLysSerLeuGluValIleThrIleLeuPheSerThrValSerAlaVal	428
Db	1024	TGGCCCAAGCGGGGAAGAGCCTGAGGTGTACCGCGCTTCACACCGCTGATGAGTCCGTG	1083
QY	429	PheGlnGlyPheAlaValCysValTyrThrHisMetAlaAspIleTrpGluValPheAsnGly	448
Db	1084	TTCACAGGCTTGGCTGTGTGTGTACCACTGACAGCATGCGAGCATCTGGAGGTTTCAACGGG	1143
QY	449	ProPheAlaHisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyGlyLysValPro	468
Db	1144	CCCTTTCGCCACCGAGTGGGCTCTCACACACAGTGGGGGCCCTAATGGGGGAGAGTCCCC	1203
QY	469	PheProArgProGlyValCysProSerLysMetThrIleGlnProGlyArgTrpProPheGly	488
Db	1204	TTCCTCTGCGCCTGGCGTGTGCCACAGCAAGATGACCGCACAGCAGGACCGCCCTTTGGC	1263
QY	489	SerThrLysAspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPhe	508
Db	1264	AGCACCAAGAGCACTCCAGATGAGGTGCACATTTGGCCGACCCACCCCTCATGTTTC	1323
QY	509	TyrProValArgProArgHisGlyArgTrpValLeuValTyrThrHisLeuAlaGlnGln	528
Db	1324	TGGCGCTGTGGCGCTGCATGCGCCCTGTCTTGTCAAGAGCCCACTGGCGCCAGCAG	1383
QY	529	LeuHisGlnIleValValAspArgValGluAlaGluAspGlyThrTyrAspValIlePhe	548
Db	1384	CTACACCAAGATGTGTGTGACCGCGTGGAGGAGAGAGATGGAGCTTACATGTCAATTTTC	1443
QY	549	LeuGlyThrAspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlyLysSerAla	568
Db	1444	CTGGGAGACTGATCGAGGCTGTGTCTCAAGTCATGTGCTTCTCAGGACAGGGGCTCAGT	1503

QY	569	GIuPProGIuGIuValValLeuGIuGIuLeuGIuValPheUlysValProThrProIleThr	588
Db	1504	GAACCTGAGGAGAGGTCTTGGAGGAGCTCCAGGTGTTAAGTGGCCACACCTATACCC	1504
QY	589	GIuMetIuIleSerValLysArgGluMetLeuTyrValGlySerArgLeuGIuValAla	606
Db	1564	GAATGAGAGTCTCTGTCCAAAGGCAATCTTACGTGGCCTCTCCGTGGGTGGCC	1628
QY	609	GIuLeuArgLeuHisGlnCysGluThrTyrGIuThrAlaCysAlaGluCysCysLeuAla	628
Db	1624	CACCTCGGGCTGCACCAATGTGAGACTTACGGACCTGCGCTGTCAGAGTCTGCTGGCC	1688
QY	629	ArgAspProTyrCysAlaTrpAspGIuAlaSerCysThrHisTyrArgProSerLeuGIu	648
Db	1684	CGGAGCCCACTACTGTGCTGGGATGGTGGCTCTGTACCACATCACGCCCACTTGGC	1744
QY	649	LysArgArgPheArgArgGlnAspIleArgHisGIuAsnProIleLeuGlnCysLeuGIu	668
Db	1744	AAGCGCGGTCTCCGCGCGAGGACATCCGCGACCGCAACCTTCTGCGAGTCTGGCG	1800
QY	669	GlnSerGIuGIuGIuGIuValValGIuLeuValAlaAlaThrMetValTyrGIuThrGIu	688
Db	1804	CAGAGCAGGAAGAAAGGACAGTGGGACTTGTGGGACCAACATGGTCTACGGCACGGAG	186
QY	689	HisAsnSerThrPheLeuGlnCysLeuProLysSerProGlnAlaAlaValArgTrpLeu	708
Db	1864	CACATATAGACCTTCTCGTAGTGCTCGCCCAAGTCTCCCAAGGCTGCTGTGGCTGGCTC	192
QY	709	LeuGlnArgProGIuAspGIuGIuProAspGlnValLysThrAspGIuArgValLeuHis	728
Db	1924	TTGCAGAGGCGGAGGAGTGGAGGGCTGACCGCTGGAAGACGGAGGAGGAGCTTGGAC	198
QY	729	ThrGluArgGIuLeuLeuPheArgArgLeuSerArgPheAspAlaGIuThrTyrThrCys	748
Db	1984	ACGGAGCGGGGGCTCTTCTCCGAGGCTTACCGCTTTCGATGGCGGACCTACACCTGC	204
QY	749	ThrThrLeuGlnHisGIuPheSerGlnThrValValArgLeuAlaLeuValIleVal	768
Db	2044	ACCACTCTGGAGACTGGCTTCTCCAGACTGTGGTCCGCTGGCTGTGGTGTATTGTG	210
QY	769	AlaSerGlnLeuAspAsnLeuPheProGIuLupLysProGIuLupProAlaArg	788
Db	2104	GCTCACAGCTGGACMACTTCTCCCTCGGAGCCAAAGCCAGAGAGCCCCACCCCGG	216
QY	789	GIuGIuLeuAlaSerThrProProLysAlaTrpTyrLysAspIleLeuGlnLeuIleGIu	808
Db	2164	GGAGGCGTGGCTTCCACCCCAACCAAGGCGCTGTACAAGGACATCTGCTACCTATTGGC	222
QY	809	PheAlaAsnLeuProArgValAspGIuTyrCysGIuArgValITrPCysArgGIuThrThr	828
Db	2224	TTCCGCCAAGCTGCCCCGGGAGTGAAGATACTGTAGGCGCGTGTGTGAGGGCACACAGC	228
QY	829	GluCysSerGIuCysPheArgSerArgSerArgGIuLysGlnAlaArgGIuLysSerTrp	848
Db	2284	GAATGCTCAGGCTGCTTCCGAGCGGAGCGCGGGGCAAGAGCCAGGCGCAAGAGCTGG	234
QY	849	AlaGIuLeuGIuLeuGIuLysLysMetLysSerArgValHisAlaGlnHisAsnArgThr	868
Db	2344	GCGAGGCTGGAGCTAGGCAAGAAATGMAAGGCCGGGTGATGCCGAGCAATCGAGCG	240
QY	869	ProArgGlnValGIuAlaThr	875
Db	2404	CCCCGGAGGTGAGGCCACAG	2424
RESULT 4			
ID	AAH47791	standard; cDNA; 2349 BP.	
XX	AAH47791;		
XX	07-JAN-2002	(first entry)	
DE	Novel human protein (NHP) encoding cDNA sequence.		


```

QY 494 ProAspGluValLeuGlnPheAlaIarGAlaHisProLeuMetPheTrpProValArgPro 513
DB 1201 CCAGATGAGGTGCTGCAGTTTGGCCAGAGCCACCCCTCATGTTCTGGCCCTGTGGCCCT 1260
QY 514 ArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleVal 533
DB 1261 CGACATGAGCGCCCTGCTTGTCAAGACCACCTGGCCAGAGCTACACAGATGCTG 1320
QY 534 ValAspArgValGlnAlaGlnAspGlyThrTrpAspValIlePheLeuGlyThrAspSer 553
DB 1321 GTGGACCGGCTGGAGCAGAGATGGAGACTAGCATGTCTATTTCTGGGGAGCTGACTCA 1380
QY 554 GlySerValLeuLysValIleAlaLeuGlnAlaGlySerAlaGlnProGluGluVal 573
DB 1381 GGGTCTGTGCTCAAAAGTCAATGCTTCCAGAGCAGGGGGGCTCAGCTAGACTGAGAAAGT 1440
QY 574 ValLeuGlnGlnLeuGlnValPheLysValProThrProIleThrGlnMetGluIleSer 593
DB 1441 GTTCTGGAGAGCTCCAGGTGGTTTAAGTCCAAACACTATCACCAAAATGAGATCTCT 1500
QY 594 ValLysArgGlnMetLeuLysValGlySerArgLeuGlyValAlaGlnLeuArgLeuHis 613
DB 1501 GTCAAAAGCAATGCTATACGTGGGCTCTCGGCTGGGTGTGGCCAGCTGGGCTGAC 1560
QY 614 GlnCysGluThrTrpGlyThrAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 633
DB 1561 CAATGTGACACTTACGGGCTGCTGTCAGAGTGTGCTGGCCGGAGCCCATCTGCT 1620
QY 634 AlaTrpAspGlyAlaSerCysThrHisTrpArgProSerLeuGlyLysArgArgPheArg 653
DB 1621 GCCTGGAGTGGCTCTGCTCCGTACCACTACCCGCCCTTGGCAGAGCCGGTTCGCC 1680
QY 654 ArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlyGlnSerGlnGln 673
DB 1681 CGCGAGGACATCCGGCAGCGCAACCTGCTGAGTGGCTGGCCAGAGCCAGAGAA 1740
QY 674 GluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGlnHisAsnSerThrPhe 693
DB 1741 GAGCAGAGTGGAGCTTGTGGCAGCCACCATGCTCTACGGCAGCAGCAGCAGCAATG 1800
QY 694 LeuGluCysLeuProLysSerProGlnAlaAlaValArgTrpLeuGlnArgProGly 713
DB 1801 CTGGGTGCTCCGCCCAAGTCTCCCAAGCTGCTGCTGGCTGCTTGGAGAGGCGCAG 1860
QY 714 AspGluGlyProAspGlnValLysThrAspGluArgValLeuHisThrGluArgGlyLeu 733
DB 1861 GATGAGGGGCTGACGACAGTGAAGAGCAGCAGCAGCTTGTGACACAGAGGGGGCTG 1920
QY 734 LeuPheArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrThrLeuGlnHis 753
DB 1921 CTGTTCGCCGAGCTTAGCGCTTGTGATCGGGGCACTACCTGACCACTCTGGAGCAT 1980
QY 754 GlyPheSerGlnThrValValArgLeuAlaLeuValValIleValAlaSerGlnLeuAsp 773
DB 1981 GGGTCTTCCGAGCTGTGTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
QY 774 AsnLeuPheProGluProGluProLysProGluGlnProAlaArgGlyGlyLeuAlaSer 793
DB 2041 AACCTGTTCCTCGGAGCCAAAGCCAGAGAGCCGCCAGCCGGGAGGCTGGCTTCC 2100
QY 794 ThrProProLysAlaIlePyrTyrLysAspIleLeuGlnLeuIleGlyPheAlaAsnLeuPro 813
DB 2101 ACCCCACCCAGAGGCTGTGTACAAAGACATCTTGACCTCATTTGGCTTGGCCACCTGCC 2160
QY 814 ArgValAspGluTrpCysGluArgValTrpCysArgGlyThrTrpGluCysSerGlyCys 833
DB 2161 CCGGTGATAGTACTTGTAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
QY 834 PheArgSerArgSerArgGlyLysGlnAlaArgGlyLysSerTrpAlaGlyLeuGlnLeu 853
DB 2221 TTCGGAGCGGAGCGGGGAGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCT 2280

```

```

QY 854 GlyLysLysMetLysSerArgValHisAlaGlnHisAsnArgThrProArgGluValGln 873
DB 2281 GGCAGAGAGATGAAGAGCCGGGTGCATGCCAGACACAAATGGAGCGCCGGAGAGTGGAG 2340
QY 874 AlaThr 875
DB 2341 GCCACG 2346

RESULT 5
AAFA90250
ID AAF90250 standard; DNA: 2340 BP.
XX
AC AAF90250:
XX
DT 06-AUG-2001 (first entry)
XX
DE Nucleotide sequence of a semaphorin polypeptide designated ZSMF-16.
XX
KW Human; semaphorin; ZSMF-16; neurite growth; neurite outgrowth;
KW T lymphocyte suppressor; cancer; stroke; brain damage; paralysis;
KW spinal injury; neurodegenerative disease; amyotrophic lateral sclerosis;
KW Alzheimer's disease; Huntington's disease; Parkinson's disease;
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW immunosuppression; autoimmune disease; insulin dependent diabetes;
KW rheumatoid arthritis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..2340
FT /tag= a
FT /product= "semaphorin"
FT sig_peptide
FT 1..66
FT /tag= b
FT /note= "Secretory signal sequence"
FT mat_peptide
FT 67..2340
FT /tag= c

WO200140278-A2.
07-JUN-2001.
06-DEC-2000; 2000WO-US3116.
06-DEC-1999; 99US-045560.
(ZYMO ) ZYMOGENETICS INC.
Holloway JL, Foley KP:
WPI: 2001-374784/39.
P-PSDB: AAB84219.

Novel human semaphorin polypeptide, ZSMF-16, useful for treating
peripheral neuropathies Alzheimer's and Huntington's disease and
polynucleotide encoding ZSMF-16 useful for detecting genetic
abnormality and cancer.
XX
PS Claim 2; Page 116-121; 124pp; English.
XX
CC The present sequence encodes a human semaphorin polypeptide, designated
CC ZSMF-16. ZSMF-16 is a neurite growth and development modulator. It also
CC enhances spinal cord and sensory neurite outgrowth and patterning, and
CC is involved in the activation and regulation of T lymphocytes suppressor.
CC ZSMF-16 is useful for detecting a genetic abnormality or cancer. ZSMF-16
CC polynucleotide probes can be used to detect 3p21 loss, trisomy,
CC duplication or translocation associated with mammary tumour tissue,
CC breast tumour, liver, small intestine, bone cancers, etc.. ZSMF-16 can
CC be used to modulate neurite growth and development and demarcate nervous
CC system structures. ZSMF-16 are also useful for regenerating and directing
CC neurite outgrowth following strokes, brain damage caused by head
CC injuries, paralysis caused by spinal injuries, and for treating
CC neurodegenerative diseases such as amyotrophic lateral sclerosis,

```

CC Alzheimer's disease, Huntington's disease, Parkinson's disease and
 CC peripheral neuropathies, or demyelinating diseases e.g., multiple
 CC sclerosis. ZSMF-16 also acts as a mediator of immunosuppression,
 CC and thus useful for diagnosing and treating autoimmune diseases such as
 CC insulin dependent diabetes, rheumatoid arthritis, and multiple
 CC sclerosis. It can also be used as an anti-inflammatory for inhibition
 CC of antigen in humoral and cellular immunity and for immunosuppression.
 CC in graft and organ transplants.

XX Sequence 2340 BP: 406 A; 748 C; 743 G; 443 T; 0 other:

Alignment Scores: 1.62e-213 Length: 2340
 Pred. No.: 4046.50 Matches: 760
 Score: 95.00% Conservative: 1
 Percent Similarity: 95.00% Mismatches: 39
 Best Local Similarity: 85.26% Indels: 2
 Query Match: 22 Gaps: 2

US-09-813-290-2 (1-875) x AAF90250 (1-2340)

QY 94 MetalaProseralaTrpAlaIleCysTrpLeuLeuGlyGlyLeuLeuLeuHisGlyGly 113
 DB 1 ATGGCCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
 QY 114 SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTyArgAsp 132
 DB 61 ACCTGTGGCCCGAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 133 -----LeuLeuSer 135
 DB 121 ATGGTCGGAACCTTCACGACCATGTGATGGAACATTTTCCAGATACCTCGCTGTC 180
 QY 136 AlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTy 155
 DB 181 GCCACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 QY 156 LeuAspGlyTyArgAspArgPheLeuGlyGlyLeuAspAlaLeuTySerLeuArg 175
 DB 241 CTAGATGATGACGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 176 LeuAspGlnAlaTrpProAspProArgGlyValLeuTrpProGlnProGlyGlnArg 195
 DB 301 CTGACACGAGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330
 QY 196 GluGluCysValArgGlyGlyArgAspProLeuThrGluCysAlaAsnPheValArgVal 215
 DB 331 -----ACAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 357
 QY 216 LeuGlnProHisAsnArgThrHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThr 235
 DB 358 CTACAGGCTGACACACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 417
 QY 236 CysAlaLeuLeuThrValGlyHisArgGlyGlnHisValLeuHisLeuGlnProGlySer 255
 DB 418 TGTGCGCTCTACACAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 477
 QY 256 ValGluSerGlyArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPhe 275
 DB 478 GTGGAAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
 QY 276 IleAspGlyLeuLeuTyThrGlyLeuThrAlaAspPheLeuGlyArgGlnAlaMetIle 295
 DB 538 ATAGACGGGAGGTGTACACGCGGCTGCTACTGCTGCGGCGGCGGCGGCGGCGGCGG 597
 QY 296 PheArgSerGlyGlyProArgProAlaLeuArgSerAspSerAspGlnSerLeuLeuHis 315
 DB 598 TTCCGAATGAGGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 657
 QY 316 AspProArgPheValMetAlaAlaArgIleProGluAsnSerAspGlnAspAsnAspLys 335
 DB 658 GACCCCGGCTTTGTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 717

QY 336 ValTyRPhePhePheSerGluThrValProSerProAspGlyGlySerAsnHisValThr 355
 DB 718 GTGACTCTCTCTCTCTGAGACCGGTCCCTCGGCCGATGTGGCTGACACCATGTCTACT 777
 QY 356 ValSerArgValGlyArgValCysValAsnAspAlaGlyGlyGlnArgValLeuValAsn 375
 DB 778 GTACAGCCCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 837
 QY 376 LysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGlyLysAla 395
 DB 838 AATGAGACACTCTCTCTCAAGCCAGGCGTGTCTGTGGTCCCGGCGGCGGCGGCGGCGG 897
 QY 396 GluThrHisPheAspGlnLeuGlnAspValPheLeuLeuTrpProLysAlaGlyLysSer 415
 DB 898 GAGACCACTTGTGACAGCTAGAGATGTGTCTCTGTGGCCCAAGGCGGGAAGAGC 957
 QY 416 LeuGluValTyArgAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
 DB 958 CTGAGGTGTAGCGGCTGTTCAGCACCGGTGAGTCCGCTTCCAGGCGCTCCCGCTCTCT 1017
 QY 436 ValTyHisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisArgAspGly 455
 DB 1018 GTGTACACATGGACAGACATCTGGAGGTTTTCACAGGCGGCGGCGGCGGCGGCGGCG 1077
 QY 456 ProGlnHisGlnTrpGlyProTyArgGlyLysValProPheProArgProGlyValCys 475
 DB 1078 CCTCACACACAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1137
 QY 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyProAsp 495
 DB 1138 CCAGACAGATGACCCGACAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1197
 QY 496 GluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgProArgHis 515
 DB 1198 GAGTGTCTGAGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1257
 QY 516 GlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleValAlaAsp 535
 DB 1258 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1317
 QY 536 ArgValGluAlaGluAspGlyThrTyArgAspValIlePheLeuGlyThrAspSerGlySer 555
 DB 1318 CGGTGTAGGACAGAGTGGAGTACCTAGATGTCTCTGCGGCGGCGGCGGCGGCGGCGG 1377
 QY 556 ValLeuLysValIleAlaLeuGlnAlaGlySerAlaGlnProGluGluValValLeu 575
 DB 1378 GTGCTCAAGTCATGCTGTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1437
 QY 576 GluGluLeuGlnValPheLysValProThrProIleThrGluMetGluIleSerValLys 595
 DB 1438 GAGAGCTCCAGGTGTTAAGTCCCAACCTATCACCGCAATGAGATCTGTCTGAAA 1497
 QY 596 ArgGlnMetLeuTyValGlySerArgLeuGlyValAlaGlnIleuArgLeuHisGlnCys 615
 DB 1498 AGCGAAATGCTATACGTGGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1557
 QY 616 GluThrTyArgIleThrAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrp 635
 DB 1558 GAGACTTACGCACTGCTGTGCGAGAGTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1617
 QY 636 AspGlyAlaSerCysThrHisTyArgProSerLeuGlyLysArgArgPheArgArgGln 655
 DB 1618 GATGTGCTCTCTGTACCACTACCGGCCCAAGCTTGGCAAGCGGCGGCTCCGCGGAG 1677
 QY 656 AspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlyGlnSerGlnGluGluValAla 675
 DB 1678 GACATCCGCGACGGACCGCTCCCTGCAAGTGTGGGCGGCGGCGGCGGCGGCGGCGG 1737
 QY 676 ValGlyLeuValAlaAlaThrMetValTyArgIleGlnHisAsnSerThrPheLeuGln 695
 DB 1738 GTGGACTTGTGGAGCCACCATGTGTACGCGGAGACCAATAGACCTTCTCTGGAG 1797
 QY 696 CysLeuProLysSerProGlnAlaAlaValArgTrpLeuGlnArgProGlyLysGlu 715

QY 256 ValGluSerGlyArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPhe 275
 Db 478 GTNGARWMSGNGMGNGMGNTGTCNCATGACARCCNMSNMGCCNTTGGCWMSACNTTY 537
 QY 276 IleAspGlyGluLeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetIle 295
 Db 538 ATHGAYGNGARATYTTATACNGAYTNACNGATTTTYYTNGGNGMGARCATGATH 597
 QY 296 PheArgSerGlyGlyProArgProAlaLeuArgSerAspSerArgPheGlnSerLeuHis 315
 Db 598 TTYMGNMSGNGGNGCCMGNGCCNTTNGMNSNGAYWSNGAYCARMSYTYNTNCA 657
 QY 316 ASPProArgPheValMetAlaAlaArgIleProGluAsnSerAspGlnAspAsnAspLys 335
 Db 658 GAYCCNMGTTYGTATGAGCNCNGNATGCCNGARAAWMSGAYCARGAYAAAYGAAR 717
 QY 336 ValTyrPhePhePheSerGluThrValProSerProAspGlyGlySerAsnHisValThr 355
 Db 718 GTNTATYTTTYYTWSNGARACNCTNCCNMGNGAYGGGWSMNAICATGTACN 777
 QY 356 ValSerArgValGlyArgValCysValAsnAspAlaGlyGlyArgValLeuValAsn 375
 Db 778 GTNWSMGNGTNGMNGNTNTGYNATYAGCGNGGNGCARMGNTNTNATNAY 837
 QY 376 LysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyAla 395
 Db 838 AARTGSMNACNTTYYTNAARGCNMGNTNTGTYWSNGTNGCCNGCNGCNGNGNGCN 897
 QY 396 GluThrHisPheAspGlnLeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSer 415
 Db 898 GARACCATTTTGATCARTYNGARAGTNTTYYTNTTGGCCNARCCNCGNANARMSN 957
 QY 416 LeuGluValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
 Db 958 YTGAGGTNTATYCCNTNTTYMSNACNGTNGTTCARAGNTTYGCCNGTNTCY 1017
 QY 436 ValTyrHisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisArgAspGly 455
 Db 1018 GTNTAYCAVATGGCNGAYATHTGGAGTNTTAAATGAGCCNTTYGCNCAYMNGAYGN 1077
 QY 456 ProGlnHisGlnTrpGlyProTyrGlyGlyLysValProPheProArgProGlyValCys 475
 Db 1078 CCNCAKATYCARGTGGGNCNTAYGGNGNARGTNCCNTTYCCNMGCCNGGNGTNG 1137
 QY 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAsp 495
 Db 1138 CCNMSNAARATGACNGCNCARCCNGMGNGCCNTTYGGNMSNACNARATATCCNGAY 1197
 QY 496 GluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgProAlaHis 515
 Db 1198 GARGTNTNTCAKRTTYCCMGNGCCATCCNTNTATGTTTGGCCNGTNGMCCMGNCAY 1257
 QY 516 GlyArgProValLeuValLysThrHisLeuAlaGlnLeuHisGlnIleValValAsp 535
 Db 1258 GGMNGCCNGTNTNTNAARACNCAVATYTCNCARCATYTCARATHTGTGTNGAY 1317
 QY 536 ArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThrAspSerGlySer 545
 Db 1318 MGNNGTNGARGCGARGAYGNAACNTATGATGTNTHTTYYTNGNACGATWMSGNMSN 1377
 QY 556 ValLeuLysValIleAlaLeuGlnAlaGlyLysSerAlaGluProGluGluValValLeu 575
 Db 1378 GTNTYNTNARGTATHTGCTNARCCNGGNGMNGNGNARCCNGARGTNGTNTN 1437
 QY 576 GluGluLeuGluValPheLysValProThrProIleThrGluMetGluIleSerValLys 595
 Db 1438 GARGATYTCAGTNTTAAATGTCNCCNACNATHTACGATGATGATHTMSNTNAR 1497
 QY 596 ArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnCys 615
 Db 1498 MGNCAATGTNTATYCTNGNMSNMGNNTNGGNTNGCNCATYTMNGNTNCAVCAKTYG 1557

QY 616 GluThrTyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrp 635
 Db 1558 GARACNTATGAGNACNCCNTGTCNGARGTGTGYTNGCMGNGAACCNTATYGCNTNG 1617
 QY 636 AspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArgArgPheArgGln 655
 Db 1618 GAYGNGCNSMNTGYACNCAVATYTMGNCNMSNYTNGGNAARNGNTTYGNGMNCAR 1677
 QY 656 AspIleArgHisGlnLysProAlaLeuGlnCysLeuGlyGlnSerGlnGluGluAla 675
 Db 1678 GAYATHTGNCAYGNAAYCCNCTNTNCAATGTCTYTTGNCARMSNARARARARGCN 1737
 QY 676 ValGlyLeuValAlaAlaThrMetValTyrGlyThrGluHisAsnSerThrPheLeuGlu 695
 Db 1738 GTNGTNTNGTNGCNCNACNATGTTATYAGNACNGARCAVAYMSNACNTTYTNGAR 1797
 QY 696 CysLeuProLysSerProGlnAlaAlaValAlaArgTrpLeuGlnArgProGlyAspGlu 715
 Db 1798 TGYTNCNCAARMSNCCNCAKCGCNGTNGTNGTNTNCAATGTCNARMGCCNGNGAYGAR 1857
 QY 716 GlyProAspGlnValLysThrAspGluArgValLeuHisThrGluArgGlyLeuLeuPhe 735
 Db 1858 GGNCCNGATYCARGTNARACNGAGARGMGNTNTNCAVACNGARMGNGNTNTTNTTY 1917
 QY 736 ArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrThrLeuGluHisGlyPhe 755
 Db 1918 MGNMGNTNMSMGNTTYGAYCGNGNACNTATYACNTGYACNACNTNNGARCAVAGTTY 1977
 QY 756 SerGlnThrValAlaArgLeuAlaLeuValIleValAlaSerGlnLeuAspAsnLeu 775
 Db 1978 WSNCAKACNCTNTGTMGTNTGTCNTNTGNTNTHTGNGCNSCARATYTGAAVAYTN 2037
 QY 776 PheProProGluProLysProGluGluProProAlaArgGlyGlyLeuAlaSerThrPro 795
 Db 2038 TTYCCNCCGARGCNAARCCNGARGARCCNCCNMGNGGNGGTYTNGCWSNACN 2097
 QY 796 ProLysAlaTrpTyrLysAspIleLeuGlnLeuIleGlyPheAlaAsnLeuProArgVal 815
 Db 2098 CCNARGCNTGTATTAARATATHTNCAKRTYNTATHTGNTTYGCNAAYTTCNMGNGT 2157
 QY 816 AspGluTyrCysGluArgValTrpCysArgGlyThrThrGluCysSerGlyCysPheArg 835
 Db 2158 GAYGATATGTGARGMGNTNTGTYMGNGSNACNACGARTGYWSNGNTGYTNGN 2217
 QY 836 SerArgSerArgGlyLysGlnAlaArgGlyLysSerTrpAlaGlyLeuGluLeuGlyLys 855
 Db 2218 WSNMGWSMNGNAGARCARCCNMGNGGNAARWSNTGGCGNTNTNARTNGNAR 2277
 QY 856 LysMetLysSerArgValHisAlaGluHisAsnArgThrProArgGluValGluAlaThr 875
 Db 2278 AARATGAARMSMNGNTNCAVCGNGARCAVAYMGNACCCNMGNGARGTNGARCCNACN 2337
 Db 2337

RESULT 7
 AAV35367 standard; cDNA; 2898 BP.
 ID AAV35367;
 AC AAV35367;
 XX 01-OCT-1998 (first entry)
 DT XX
 DE Human semaphorin encoding cDNA.
 KW Human; semaphorin; diagnosis; nerve disease; nerve growth inhibitor;
 KW neurological disease; atopic skin inflammation; autoimmune disease;
 KW pain; ds.
 OS Homo sapiens.
 OS XX
 OS XX
 OS XX
 FH key Location/Qualifiers
 FT 370..2697
 FT CDS /*tag= a
 FT /product= "semaphorin"
 FT XX

PN W09822504-A1.
 XX 28-MAY-1998.
 XX 12-NOV-1997; 97WO-JP04111.
 XX 15-NOV-1996; 96JP-0321068.
 XX (SUNKU) SUMITOMO PHARM CO LTD.
 XX Furuyama T, Inagaki S;
 XX MPI: 1998-312416/27.
 XX P-PSDB: AAM63748.
 DR
 XX
 XX Gene encoding new semaphorin nerve growth inhibitor - useful in
 PT diagnosis, treatment and study of neurological diseases
 XX
 XX Claim 2: Page 37-39; 49pp; Japanese.
 CC The present sequence encodes human semaphorin, a nerve growth
 CC inhibitor. The semaphorin protein, and gene encoding the protein,
 CC and their derivatives, are used in the diagnosis, treatment and
 CC study of neurological disorders such as atopic skin inflammation,
 CC autoimmune diseases and pain.
 XX
 SQ Sequence 2898 BP; 794 A; 695 C; 748 G; 660 T; 1 other;
 Alignment Scores:
 Pred. No.: 3,966-104 Length: 2898
 Score: 2056.00 Matches: 418
 Percent Similarity: 63.75% Conservative: 136
 Best Local Similarity: 48.10% Mismatches: 264
 Query Match: 43.32% Indels: 51
 DB: 19 Gaps: 14
 US-09-813-290-2 (1-875) x AAV35367 (1-2898)
 QY 36 ArgGlnGlyPro-----SerLeuLeuSerSerAlaProLeuProAla 50
 Db 126 AGACCGCGTCCCTGGACGCGTCTTCTCCGCGCTTAAACCCGCGCCAGAGACAGA 135
 QY 51 GlnAsp---TyrValGluProLeuProTyrTyrTrpProGlyGlySerArgAlaAsn 69
 Db 186 AAGCGTTAGCGGATCCAAATATATGCGCGCAATGSCACTGGAAATGATTTCTGAT 245
 QY 70 TyrAsn-----ArgArgProAlaGlyProGluGly----- 80
 Db 246 GACAAACCCCTTCTGTTGTGACAAAGCCTGTGCGCCGACAGTTGCCCGAGGAGAGTAC 305
 QY 81 -----SerAlaGlyArgArg-GlnArgCysProGlnPhePr 92
 Db 306 TAAATAAACTCAATCTGTCTTAAAGTGTGCTGACAGGCGCCAGAGAGAGACCCAGCAAG 365
 QY 92 oSerMetAlaProSerAlaTrpAlaIleCysTrpLeuLeuGlyLeuLeuHisG1 112
 Db 366 CACATGTGACCGCGCGGACACATCCCTCACCCTGCTGTGGGCTGACACATAAAGA 425
 QY 112 yGlySerSerGlyProSerProGlyProSerValProArgArgLeuSerTyrArgAs 132
 Db 426 CTGACCCCGAGCTACTCCGCGGAACCCCTCTACCCGCTAGCGCTGTCACATAAAGA 485
 QY 132 PheLeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnLeuG1 152
 Db 486 ACTTTTGAACCTGAAATAGGACTTCAATATTTCAAAAGCCCTTGATTTCTTGATCTCA 545
 QY 152 nAlaMetTyrLeuAspGlyTyrArgAspArgLeuPheLeuGlyLeuAspAlaLeuTy 172
 Db 546 TACAATCTCTGATGATAGATACAGAGAGCGCTTTTGTGGAGAGAGACCTTGCTCA 605
 QY 172 rSerLeuArgLeuAspGlnAlaTrpProAspProArgGlyValLeuLeuTrpProProGlnPr 192
 Db 606 TTCCTGAACTTGAACGAGTCAAGTACGCGCTACAGAGATATACCTGCGCCAGACAGC 665

QY 192 oGlyGlnArgGluGluCysValArgLysGlyArgAspProLeuThrGluCysAlaAsnPh 212
 Db 666 AGTAAAGGTAGAAAGAAATGCAATGAAAGCAAAAGAC---GCAATATGCTGCGCAATTA 722
 QY 212 eValArgValLeuGlnProHisAsnArgThrHisLeuAlaCysGlyThrGlyAlaPh 232
 Db 723 TATCCGGGTTTGCATTCACCTACCAACAGACACACTTCTGACCTGCTACTGAGCTTT 782
 QY 232 eGlnProThrCysAlaLeuLeuThrValGlyHisArgGlyGluHisValLeu---HisLe 251
 Db 783 TGATCCACACTGTGCTTCTATCAGAGCGGCGACCAATTCAGAGAAACCCCTGTTTCACT 842
 QY 251 uGluProGlySerValGluSerGlyArgGlyArgCysProHisGluProSerArgProPh 271
 Db 843 GGAGTACACACAGATCTGAGAGAGAGAGGAGCAATGCTTTTCCATCCCAACCTCCCTT 902
 QY 271 eAlaSerThrPheLeuAspGlyLeuLeuTyrThrGlyLeuThrAlaAspPheLeuGlyTr 291
 Db 903 TGTGTCCAGCGTAGTTGGGAAATGAGCTGTTGCTGGACCTTACAGTACTATTGGGGCAG 962
 QY 291 gGluAlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAsp---SerAs 310
 Db 963 AGACTGGCGATCTTCCGACGATGGGGAAGTTAGGCCAATATTCGCATGAGCATGACGA 1022
 QY 310 pGlnSerLeuLeuHisAspProArgPheValMetAlaAlaArgIleProGluAsnSerAs 330
 Db 1023 TGAGCGCGCTCCGTAAGAAACCAAAATTTGAGGTTCTATGATGATTCCTGATACGAGA 1082
 QY 330 pGlnAspAsnAspLysValTyrTrpPhePheSerGluThrValProSerProAspGlyG1 350
 Db 1083 CCGAGATGACAAACAAATATCTTTTCTTCTGAGAGCGCGCTGAGCGGAGCAACA 1142
 QY 350 ySerAsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGlyGly1 370
 Db 1143 CGCC---CACAGATCTTACACCCGAGTGGGCGGCTGCTGCAATGATGAGAGAGACA 1199
 QY 370 nArgValLeuValAsnLysTrpSerThrPheLeuLysAlaArgValCysSerValPr 390
 Db 1200 GAGAACTCCGTGAGAACAAAGTGGAGACCTTCTTAAAGCGCGCGTGTGCTCAGTGCC 1259
 QY 390 oGlyProGlyGlyAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeuTrPr 410
 Db 1260 GCGAATGAAATGGAATCGACATCTTGTGACGACTAGAGGATGTTGTTTCTCTCCGAC 1319
 QY 410 oLysAlaGlyLysSerLeuGluValTyrAlaLeuPheSerThrValSerAlaValPheG1 430
 Db 1320 CAGAGATCCTTAAGATTCAGATGATTTGGACTGTTTAACTTACCAGCAATATATTAG 1379
 QY 430 nGlyPheAlaValCysValTyrHisMetAlaAspIleTrpGluValPheAsnGlyProPh 450
 Db 1380 AGCCATCTGTATGTGTGTATCACAATGATCCGAGAACCTTAAATGAGCCCATTA 1439
 QY 450 eAlaHisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyLysValProPhePr 470
 Db 1440 TGCATATAAAGAGCCCTGATATCCACTGTGCTACTATATGAAAGAAAGTCCCTTACC 1499
 QY 470 oArgProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySerTh 490
 Db 1500 AAGGCTGTGTTCTGTGCGACCAAAAGTAAAC-----GAGAGCAGATGTGGAACAC 1550
 QY 490 rLysAspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrPr 510
 Db 1551 CAAGATTAACCCCGATCCGCGCATCCGCTTGCAAGAGATGATCTCTTAAGTATACAGCC 1610
 QY 510 oValArgProArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnGlnGlnLeuH1 530
 Db 1611 CATAAACCTGTCTATATAAAACCAATACCTGTAAAAAACAGATGGAATAATCAACCTAG 1670
 QY 530 sGlnIleValValAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuG1 550
 Db 1671 GCAACTTGCGCTGATCGGGTGAAGCGAGGAGATGCCAGTATGACGCTTATTTATTGG 1730

QY 3 CysAlaLeuAlaGlyLysValPheProMetClySerTrp-----Pro 16
 DB 241 TCGTGTCTGCTCTTCAGCATGTTTCTCATGATATAGCTGGCGGAGACAGTACAGTACGCA 300
 QY 17 ValTrpHisLysSerLeuHis---TrpAlaAsnLysValGluGluAlaAlaGly 35
 DB 301 CAGCCCTCTCTCTACCTCCGCTGCGGTGAC-----GGCGAC 339
 QY 36 ArgGlnGlyProSerLeuLeuLeuSerSer-AlaProLeuProAlaGlnAspTrp----- 53
 DB 340 AGCAGCAGCCCGGACCGTCTCTCAAGACCGCTCTCT-----TGGACGGT 384
 QY 54 -----ValGlnPro 56
 DB 385 CTCTGCTCTCGCGCTTCTAACCACCGGCGCAAGACAGAAAGCTTAGCGGATCCAAAT 444
 QY 56 OleuProGlyLysTrpTrpProGlyLysSerArgAlaAsnTrpAsn----- 71
 DB 445 ATTGGCCCGCAATGGCAGCTGGGAATGGTATTTCTGATGACACCCCTCTGTTGTG 504
 QY 72 -----ArgArgProAlaGlyPro-GluGly----- 79
 DB 505 ACMAAGCCTGTGCGCCCGCAGTGGCCCTGGAGGGAAGTAACTAAGTAAACTCAATCTG 564
 QY 80 -----GlySerAlaGlyArgArgGlnArgCysProGlnPheProSer 93
 DB 565 TCTTAAGTGTGGCTGCGAGGGCCAGAGAGAGACCCAGCAGC-----ACC 609
 QY 94 MetAlaProSerAlaTrpAlaLeuLysTrpLeuLeuGlyGlyLeuLeuLeuHisGlyGly 113
 DB 610 ATGGACACCGCGCGAGACATCTCTACACCTTGCGTCTGGGGTCTACCTGCTGAACTCTGG 669
 QY 114 SerSerGlyProSerProGlyProSerValProArgLeuArgLeuSerTrpArgAspLeu 133
 DB 670 ACCCGAGGCTACTGCGCGAAGCCCTCTTACCCAGGCTGAGCTGACATTAAGAACTT 729
 QY 134 LeuSerAlaAsnArgSerAlaLeuPheLeuGlyProGlnGlySerLeuLeuGlnAla 153
 DB 730 TTGGAATGAAATAGACATCATATTTCAAGCCCTTGATTTCTTATGATCTCCATACA 789
 QY 154 MetTrpLeuAspGlyTrpArgAspArgLeuPheLeuGlyGlyLeuAlaLeuTrpSer 173
 DB 790 ATGCTGTGGATGATGATCAAGAACGCTCTTGTGGAGCAGAGACCTGTCTATTCC 849
 QY 174 LeuArgLeuAspGlnAlaTrpProAspProArgGluValLeuTrpProGlnProGly 193
 DB 850 CTGAACCTTGGAAAGAGTACAGTACGCGCTACAGAGATATCTGCGCAGCAGCAGCTA 909
 QY 194 GlnArgGluGlyCysValArgGlyLysArgAspProLeuTrpGlnCysAlaAsnPheVal 213
 DB 910 AAGGTAGAAATGATGATATGAAAGAAAGAAC---GCAAAATGATGTGCCAATTAATTC 966
 QY 214 ArgValLeuGlnProHisAsnArgPheHisLeuLeuAlaCysGlyTrpGlyAlaPheGln 233
 DB 967 CGGCTTTTGCACTACATACAGACAGCAGCCTTCTGACCTGTGCTACTGAGGACTTTTAT 1026
 QY 234 ProThrCysAlaLeuLeuTrpValGlyHisArgGlyHisValLeu---HisLeuGln 252
 DB 1027 CCACACGTGCTTCAATCAGAGTGGGAGCAGCATTCAGAGAAACCCCTGTTCACCTGAG 1086
 QY 253 ProGlySerValGlySerGlyArgGlyArgCysProHisGlnProSerArgProPheAla 272
 DB 1087 TCACACAGATCTGAGAGAGAGAGGAGAGATGCTTTTGGCCCACTCTCTCTGTG 1146
 QY 273 SerTrpPheLeuAspGlyGluLeuTrpTrpGlyLeuThrAlaAspPheLeuGlyArgGlu 292
 DB 1147 TCACACGTAGTGGAGATGAGCTGTTGCGTGGAGCTACAGATGATGATGGGAGAGAC 1206
 QY 293 AlaMetLeuPheArgSerGlyGlyProArgProAlaLeuArgSerAsp---SerAspGln 1211
 DB 1207 TCGGCGATCTTCCGACAGCATGGGAGAGTAGGCCATATTCGACACTGAGCATGACGATGAG 1266

QY 312 SerLeuLeuHisAspProArgPheValMetAlaAlaArgLeuProGluAsnSerAspGln 331
 DB 1267 CGGCTCTGAAAGAACCAAAATTTGTAGTTCATATGATTTCTGTATTAACGAAGACGGA 1326
 QY 332 AspAsnAspLysValTrpPhePhePheSerGlnTrpValProSerProAspGlySer 351
 DB 1327 GATGACACAAACAAATGCTCTTTCTTCTTCTAGAGAGCGCTGAGGCGGCAACAGGCC 1386
 QY 352 AsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGlyGlnArg 371
 DB 1387 ---CACAGATTTACACCCAGATGGGCGGCTGTGCTGATGATGACATGGAGAGACAGA 1443
 QY 372 ValLeuValAsnLysTrpSerTrpPheLeuLysAlaArgLeuValCysSerValProGly 391
 DB 1444 ATCTGTGTAACAAGTGGAGGAGCTTCTTAAAGCGGCGGTGGTCTGACGCGCGGA 1503
 QY 392 ProGlyValAlaGlnTrpHisPheAspGlnLeuGlnAspValPheLeuLeuTrpLys 411
 DB 1504 ATGAATGGAATCGACACATGCTTGAAGAGACTAGAGATGTGTTTACTGCCGACCA 1563
 QY 412 AlaGlyLysSerLeuGluValTrpAlaLeuPheSerThrValSerAlaValPheGlnGly 431
 DB 1564 GATCTTAAGATCCAGTGAATTTGAGCTGTTAATATACACAGCAATATATTAGAGGC 1623
 QY 432 PheAlaValCysValTrpHisMetAlaAspLeuTrpGluValPheAsnGlyProPheAla 451
 DB 1624 CATGCTGTATGTGTATATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1683
 QY 452 HisArgAspGlyProGlnHisGlnTrpGlyProGlyLysValProPheProArg 471
 DB 1684 CATAAAGAGGCGCTCAATACCTGCTGATATGAGAGAAATCCCTTACCCCAAG 1743
 QY 472 ProGlyValCysProSerLysMetTrpAlaGlnProGlyArgProPheGlySerTrpLys 491
 DB 1744 CTGCTGTCTGTGCGCGCAAGTAAC-----GAGGCAAGATGGAACCCACCAAA 1794
 QY 492 AspTrpProAspGlyValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProVal 511
 DB 1795 GATTTACCCGATGAGCCATCCGCTGCGCAAGGATGATGATGATGATGATGATGATGAT 1854
 QY 512 ArgProArgHisGlyArgProValLeuValLysTrpHisLeuAlaGlnLeuHisGln 531
 DB 1855 AAACCTGTTCAATAAACAATACCTGCTGTAACAGATGAAATACAACTGAGGCA 1914
 QY 532 LeuValValAspArgValGluAlaGlnAspGlyTrpTrpArgValPheLeuGlyTrp 551
 DB 1915 CTGGCGTGGATGCGGATGGAAGCGGAGATGGCCAGTATGAGCTGTTATTTTGGACA 1974
 QY 552 AspSerGlySerValLeuLysValLeuAlaLeuGlnAlaGlySerAlaGlnProGlu 571
 DB 1975 GACACAGGATTTGCTGAAAGTATCAATTTACAAACCAAGAAACAGAGTGGAG 2034
 QY 572 GluValValLeuGlnGluLeuGlnValPheLysValProThrProIleThrGlnMetGlu 591
 DB 2035 GAAGTCATTTAGAGAACTTCAATATTCAGAGATCCAGGCCCTTATTTCTATAGGAA 2094
 QY 592 IleSerValLysArgGlnMetLeuTrpValGlySerArgLeuGlyValAlaGlnLeuArg 611
 DB 2095 ATTCTTCAAGAGACAAACAGCTTTACATTTGATGATGATGATGATGATGATGATGATGAT 2154
 QY 612 LeuHisGlnCysGluTrpTrpGlyTrpHisLysAlaCysLysCysLeuAlaArgAspPro 631
 DB 2155 TTTCATGCTGCGAGCATATGATGAGCAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2214
 QY 632 TyrCysAlaTrpAspGlyAlaSerCysThrHisTrpArgPro-----SerLeuGlyLys 649
 DB 2215 TACTGTGCTGTGGATGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274
 QY 650 ArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLysGlyGln 669
 DB 2275 AGGAGGTTCCGAGGAGAGAGGATGCGATGCGCAAGCGCGCAAGAGTGTGGAGAG 2334
 QY 670 SerGlnGluGlnAlaValGlyLeuValAlaAlaLeuMetValTrpGlyTrpHisGlnHis 689


```
QY 80 -----GlySerAlaGluArgGlnArgCysProGlnPheProSer 93
Db 565 TCCTAAAGTGTGGCTGACGGGGCCAGAGAGAGCAGCAGCC-----ACC 609
QY 94 MetLapProSerAlaIlePheLeuLeuGlyLeuLeuLeuHisGly 113
Db 610 ATGGACCCGGCCGAGACATCTCAGCTTGGCTGGGCTGACACTGTGAACTGG 669
QY 114 SerSerGlyProSerProGlyProSerValProArgLeuLeuSerTyrArgAspLeu 133
Db 670 ACCCAGGTCCTGGGAGAACCCCTCCACCCAGGCTACCGCTGTGCATTAAGAACT 729
QY 134 LeuSerAlaAsnArgSerAlaIlePheLeuGlyProGlnGlySerLeuAsnGlnAla 153
Db 730 TTCGAACTGATGACTTCATATATTCAAGCCCTTGGATTTCCTGATCTCATACA 789
QY 154 MetTyrLeuAspGluTyrArgAspArgLeuPheLeuGlyLeuLeuPheLeuTyrSer 173
Db 790 ATGCTGCTGATGATGATCAAGAAAGGCTCTTGTGGAGGACAGACCTTGTCTATTC 849
QY 174 LeuArgLeuAspGlnAlaIlePheProArgGluValLeuTyrProGlnProGly 193
Db 850 CTGAACCTTGGACAGAGAGAGAGGCTACAGAGATATATCTGCCGAGACAGCACTA 909
QY 194 GlnArgGluLeuCysValArgGlyArgAspProLeuThrGlyCysAlaAsnPheVal 213
Db 910 AACGTAAAGAAATGCAATTAATGAAGAAAGAAC---GCAATGAGTGTGCCAATATATC 966
QY 214 ArgValLeuGlnProHisAsnArgThrHisLeuLeuAlaCysGlyThrGlyAlaPheGln 233
Db 967 CGGGTTTGTGATCATCAACAGAGACACCTTGTGCTGCTGTCTGTCTGTCTGTGAT 1026
QY 234 ProThrCysAlaLeuIleThrValGlyHisArgGlyGlnHisValLeu---HisLeuGln 252
Db 1027 CCACACTGTGCTTCATCAAGAGTGGGACCATTCAGACAGAACCCCTGTTCACCTGGAG 1086
QY 253 ProGlySerValGluSerGlyArgGlyArgCysProHisLeuProSerArgProPheAla 272
Db 1087 TCACACAGATCTAGAGAGAGAGAGAGAGATGCTTTTGAACCCCACTCTCTCTGTGTG 1146
QY 273 SerThrPheIleAspGlyLeuLeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGln 292
Db 1147 TCCACCTAGTGTGGATAGCTGTGTGTGCTGACCTGATCATATTTGGGCGAGAGAC 1206
QY 293 AlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAsp---SerAspGln 311
Db 1207 TCGGCACTCTCGAGACATGGGGAAGTTAGGCCATATTCGCACTGACATGACATGAG 1286
QY 312 SerLeuLeuHisAspProArgPheValMetAlaAlaArgIleProGluAsnSerAspGln 331
Db 1267 CGGCTCTGAAAGAACCAAAATTTGATGCTCATATATGCTCTCATTAACGAAAGACGA 1326
QY 332 AspAsnAspLysValTyrPhePhePheSerGluThrValProSerProAspGlySer 351
Db 1327 GATGACAAACAAATGATCTTTCTTACTGAGAAAGCCCTGAGAGCGGAGAACACGCC 1386
QY 352 AsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGlyGlnArg 371
Db 1387 ---CACACAGATCTTACACCCGAGTGGGCGCTGTGCTGTAATGACATGGAGAGACAGA 1443
QY 372 ValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGly 391
Db 1444 ATCCGTGCTGAACAAGTGAAGCACTTCCCTTAAAGCGCGCTGTGTGCTGATGCGCGGA 1503
QY 392 ProGlyGlyAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeuTrpProLys 411
Db 1504 ATGAATGGAATTCGACATACTTGTGACGAACCTAGAGAGATGTTTACTGCGGACCA 1563
QY 412 AlaGlyLysSerLeuGluValTyrAlaLeuPheSerThrValSerAlaValPheGlnGly 431
Db 1564 GATCTTAAGATCCAGTATATTTGACTGTTTAATACACGACATATATATTAGAGGCG 1623
QY 432 PheAlaValCysValTyrHisMetAlaAspIleTyrGluValPheAsnGlyProPheAla 451
1624 CATGCTGATGTGTATCATCATGATCAAGTATCCGGGAAGCCTTAAAGGCCATATGCT 1683
QY 452 HisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyGlyValProPheProArg 471
Db 1684 CATAAAGAGGCCCTGATATACACCTGCTACTATATGAAGAAAGTCCCTACCCCAAGG 1743
QY 472 ProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLys 491
Db 1744 CCGTGTCTCTGCGCAGCAAGTAAC-----GGAGCAGATATGACCAACCA 1794
QY 492 AspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProVal 511
Db 1795 AGATTACCCGATGACGCCATCCGCTGCGAAGATGCTCTTCAATGTATACCCCATTA 1854
QY 512 ArgProArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnLeuHisGln 531
Db 1855 AAACCTGTCTATAAAACCAATCTGTGTAATAACATGAAATACCACTGAGCGAA 1914
QY 532 IleValAlaAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThr 551
Db 1915 CTGCCCGTGAATCGGGTGAAGCGAGAGATGCGCAGATGACGTCTTATTTATTTGGCACA 1974
QY 552 AspSerGlySerVal---LeuLysValIleAlaLeuGlnAlaGlySerAlaGluPro 570
Db 1975 GACACAGGAATTTGTGCTGCTGAAGTATATCAATTTACACCAAGAAAGAGTGGATG 2034
QY 571 GluGluValValLeuGluGluLeuGlnValPheLysValProThrProIleThrGluMet 590
Db 2035 GAGGAAGTCTATTTAGAGAACTTCAAAATATTCAAAGATGCCAGCCCTATCATTTGATG 2094
QY 591 GluIleSerValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeu 610
Db 2095 GAATTTCTTCAAGACACACACACCTTTACATGTGATGACCTCTGCTGTGGCAAGTCT 2154
QY 611 ArgLeuHisGlnCysGluThrTyrThrAlaCysAlaGluCysCysLeuAlaArgAsp 630
Db 2155 AGATTCATCATCTCGACATGATGATGCGACGTGCTGTCTACTGTCTGCTGCGTGGAGAC 2214
QY 631 ProThrCysAlaIlePheAspGlyAlaSerCysThrHisTyrArgProSerLeuGly----- 648
Db 2215 CCGTACTGTCTCTGGATGGATCATATCTGCTCCAGGATCAACCAACAGGTGGCACAGAA 2274
QY 649 LysArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGly 668
Db 2275 AAGAGGAGTTCGCGAGCGAGACGTTGCGCATGGCCACCCGCCAACAGTCTTTGGA 2334
QY 669 GlnSerGlnGluGluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGlu 688
Db 2335 CAGCAATTTGTTGGAACGCTTGGACAGAGACTGAAGAGAGGCTGTATGCGCATAGAG 2394
QY 689 HisAsnSerThrPheLeuGlnCysLeuProLysSerProGlnIleAlaValArgTrpLeu 708
Db 2395 AGCAACAGTACTGTGTGGATGATCACCCCGCTCTCATCAACACCAAAAGCATGTGTTGT 2454
QY 709 LeuGlnArgProGlyLysAspGluGlyProAspGlnValLysThrAspLysArgValLeuHis 728
Db 2455 CTACAGAAAGGACGCGACGTAAGAAAGAAAGAGAGATGACAGCGATGACAGTGTGCAG 2514
QY 729 ThrGluArgGlyLeuLeuPheArgValGlyLeuSerArgPheAspAlaGlyThrTyrCys 748
Db 2515 ATGGACTTGGGCTGTCTCTCAAGATGACCAAGTCAAGATGACAGGACCTATTTTGTGC 2574
QY 749 ThrThrLeuGluHisGlyLysSerGlnThrValAlaArgLeuAlaLeuValIleVal 768
Db 2575 CAGACAGTGAACACATATTTGTGCCATCTACTGTGCGTAATAATCACTTGGAGGTGCGAA 2634
QY 769 AlaSerGlnLeuAspAsnLeuPheProProGluProLysProGluGlu----- 784
Db 2635 GAGCATTAAGTGGAGGCGCTTTTCATAGACCATGAGAGAGAAAGACATCAACAAGATG 2694
QY 785 -----ProProAlaArgGlyGlyLeuAlaSerThrProProLysAlaIleTrpTyrLysAsp 802
```

Db 2695 CCCTGCCCCCTTAAAGCGGTATGCTCAGGACGACAAACCG-----TGCTACAGGAA 2748
 QY 803 ILeuGlnLeuIleGlyPheAlaAsn---LeuProArgValaIspGlyTyrCysGluArg 821
 Db 2749 TTCCTTGACGATGATGGCTACAGCAGCAAGTCCAGAGATGGAAGATGCTGCAAAAG 2808
 QY 822 ValTrrCysArgGlyThrThiGluCysSerGlyCysPheArgSerArgSerArgIlys 841
 Db 2809 GTGTGGGTGTACAGAT-----AAGACAGAGAAAAAGCTTAAA 2844
 QY 842 GlnAlaArgGlyLysSerTrpAlaGlyLeuGlyLysLysMetLysSerArgVal 861
 Db 2845 ATGCTCCCTCCCAAGTGAAGTATGCCAACCCCGAAGAAAGAGCTTCCTTAAA--- 2901
 QY 862 HisAlaGlnHisAsnArgThrProArg 870
 Db 2902 ---GCTGAGCACTTCGCGCCAGG 2925
 RESULT 10
 AAH47049
 ID AAH47049 standard; DNA; 2709 BP.
 AAH47049;
 29-OCT-2001 (first entry)
 Semaphorin D cDNA sequence.
 Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen
 24p3; proliferation; maspin; cancer; cytostatic; gene therapy; ss.
 Homo sapiens.
 WO200155455-A2.
 02-AUG-2001.
 31-JAN-2001; 2001WO-US03161.
 31-JAN-2000; 2000US-0179191.
 (MILL-) MILLENNIUM PHARM INC.
 (JINS/) JIN S.
 Jin S;
 WPI; 2001-488799/53.
 Determining if a compound modulates the drug resistance of a cell,
 comprises determining the expression or activity level of a resistance
 sequence in a cell in the presence of the test compound
 Example 1; Fig 1A-B; 79pp; English.
 The invention relates to a method of determining whether a test compound
 modulates the drug resistance of a cell that comprises determining the
 expression or activity level of resistance genes (e.g. semaphorin D, B94,
 mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
 the test compound, and comparing its expression or activity level in a
 cell without the test compound. The drug resistant sequences are useful
 in identifying drug resistant cells, in screening methods directed to the
 identification of compounds that can modulate the drug resistance of a
 cell type or multiple cell types. An isolated resistance protein can be
 used as an immunogen to generate antibodies that bind the resistance
 protein. Resistance nucleic acids may be inserted into vectors and used
 as gene therapy vectors. An anti-resistance protein antibody may be used
 to isolate a resistance protein, or facilitate the purification of
 natural resistance protein from cells and of recombinantly produced
 resistance protein expressed in host cells. The methods are useful for
 treating a subject having a disorder, such as a drug-resistance cancer,
 characterized by aberrant resistance sequence expression or activity by
 administering to the subject a resistance modulator. The present sequence
 represents a semaphorin cDNA sequence, whose expression was increased in

CC drug resistant EMT6 tumours.
 XX Sequence 2709 BP; 831 A; 560 C; 620 G; 698 T; 0 other;
 Alignment Scores:
 Pred. No.: 5.7e-92 Length: 2709
 Score: 1834.00 Matches: 378
 Percent Similarity: 62.19% Conservative: 150
 Best Local Similarity: 44.52% Mismatches: 262
 Query Match: 38.64% Indels: 60
 DB: 22 Gaps: 16
 US-09-813-290-2 (1-875) x AAH47049 (1-2709)
 QY 43 LeuSerSerAlaProLeuProAlaGlnAspTrpValGluProLeuProTyrLysTrp 62
 Db 87 CTCGGCTGTTCCTCCCATGTCTCAGCCAGTCT---ATTCCAGATGTGTTGAAGTCTCTCG 143
 QY 63 ProGlyLysArgAlaAsnTyrAsnArgArgProAlaGlyProGlyLysSerAla 82
 Db 144 CCG-----GACAAATACA 155
 QY 83 GlyArgArgGlnArgCysProGlnPheProSerMetAlaProSerAlaTrpAla----- 100
 Db 156 GGAAAGAGACTAAAGACACCAAGGAGCTTACA---GCGTCTGACAGATGGCTGGTTAA 212
 QY 101 -----IleCysTrpLeuLeuGlyLysLeuLeuHisGlyLysSerArgProSer 118
 Db 213 CTAGGATTTGCTCTCT---CTTTCTGGGAGATATTACTTACAGCAAGCAACATTCAGAAAT 271
 QY 119 ProGlyProSerValProAlaGlnLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArg 138
 Db 272 GGAAGAAACAAATGTGCCAAGGCTGAATATCTTACAAAGAAATGTGGATCCAACAAAT 331
 QY 139 SerAlaIlePheLeuLeuProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGlu 158
 Db 332 GTGATCATCTTCAATGAGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTTTGGATGAG 391
 QY 159 TyrArgAspArgLeuPheLeuGlyLysLeuAspAlaLeuTyrSerLeuArgLeuAspGln 178
 Db 392 GAAAGGAGTGGCTGTATGTGGAGCAAGATCACATATTATTCATTCAGAGTAACTTAAG 451
 QY 179 AlaTrpProAspProAlaGlyValLeuTyrProProGlnProGlyGlnArgGluGluCys 198
 Db 452 ATC---AAGATTTTCAAAAGATTTGCTGGCCAGATATCTTACACCAAGAAAGATGAATGC 508
 QY 199 ValArgLysGlyArgAspProLeuThrGluCysAlaAsnPheValArgValLeuGlnPro 218
 Db 509 AAGTGGCTGGAAGAAGACATCTGAAAGAAATGTCTTAATTCATCAGAGTAACTTAAGCA 568
 QY 219 HisAsnArgThrHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeu 238
 Db 569 TATATACAGACTCACTGTACGCTGTGGAACGGGCTTTTCAATTCATTAATTCAGCTAC 628
 QY 239 IleThrValGlnHisArgGlyGlu---HisValLeuHisLeuGlnProGlySerValGlu 257
 Db 629 ATTCAAAATTTGACATTCATCTGAGAGCAATATTTTAAAGCTGGGAATCTCAATTTTGA 688
 QY 258 SerGlyArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPheIleAsp 277
 Db 689 AAGCGCGCTGGGAAGAGATCCATGACCTTAAGCTGACAGAGATCCCTTTAATAGAT 748
 QY 278 GlyLeuLeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetIlePheArg 297
 Db 749 GAGAAATATGACTGTGAACTGACAGCTGATTTTATGGGCGAGACTTGTATCTTCCGA 808
 QY 298 SerGlyArgProArgProAlaLeuArgSerAspSer---AspGlnSerLeuLeuHisAsp 316
 Db 809 ACTCTTGGGACCAACCCCAATAGAGAGACAGACATGATTTCCAGGTGGCTCAATATAT 868
 QY 317 ProArgPheValMetAlaAlaArgIleProGluAsnSerArgIleAspAsnAspLysVal 336
 Db 869 CCAAGATTCAATTAGTGGCCACCTCATCTCAGAGACTGACAAATCTCGAAGATGACAAAGTA 928

QY	337	tyrPheRhePheSerGIuThrIvalProSerProAspGly---	GlySerAsnHisValThr	355
Db	929	TACTTTTTCCTCCGTGAATAAT-----GCATGATGGGAAACACCTCGGAATAACT		982
QY	356	ValSerArgValGlyArgValCysValAsnAspAlaGlyGlyInrValLeuValAsn		375
Db	983	CACGCTACAATAGGTCAGATATGCAAAATGACTTTTGGAGGCGACAGAAAGTCTGGTGAAT	.1042	
QY	376	LysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyAla	395	
Db	1043	AAATGGACAAATCCCAAGGCGCTGATTTGGCTCAGTGGCAGGTCCAAATGGCATTT	1102	
QY	396	GIUThrHisPheAspGlnLeuGlnAspValPheLeuLeuTrpProLysAlaLysSer	415	
Db	1103	GACACTACTTTTGATGAACCTGCAAGATGATTTCTCTAATGAACTTTAAATATCTTAATAAT		1162
QY	416	LeuGlnValTyralaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys	435	
Db	1163	CCAGTTGATATGGAGTGTTTACAGCTCCAGTAAACATTTTTCAAAGGATCAGCCCTGAT	1222	
QY	436	ValTyrlHisMetAlaAspRileTrpGlnValPheAsnGlyTrpPheAlaHisArgAspGly	455	
Db	1223	ATGTAAGACATGAGTGTATGGAGAGGGTCTCTTGGTCATATCCCAACAGGGATGGA	1282	
QY	456	ProGlnHisGlnTrpGlyProTyrglyGlyLysValProPheProArgProGlyAlaCys	475	
Db	1283	CCCAACTATCAATGGGTCCTTATCAAGAGAAAGTCCCTATTCACAGGCGACAGACTGT	1342	
QY	476	ProSerLysMetThrIaGlnProGlyArgProPheGlySerThrLysAspTryProAsp	495	
Db	1343	CCCAAGCAAAACATTTGGT-----GGTTTGGACTTCACAAAGACCTTCCTAT	1390	
QY	496	GIUValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgProArgHis	515	
Db	1391	GATGTATTAACCTTTGGCAAAAGTCAATCCAGGCATGTAAATCCCACTGTTCTATGAAC	1450	
QY	516	GlyArgProValLeuValLysThrHisLeuAlaGlnLeuHisGlnLeuValAsp	535	
Db	1451	AAATGCCCAATAGTGATCAAAAGCGATGTAAATATCAATTTTACCAAAATGTCTGTATAC	1510	
QY	536	ArgValGlnAlaGlnAspArgLysThrTyArgAspValIlePheLeuGlyThrAspSerLys	555	
Db	1511	CGATGTGATCGCAAAAGATGCAACATGATGATGTATTTATGTCGAAACAGATCTGGGAC	1570	
QY	556	ValLeuLysValIleAlaLeuGlnAlaGlyGlySerAlaGlnProGlnGlnValAlaLeu	575	
Db	1571	GTCTGTAAAGTAGTTTCAATTCCTTAAGAGACCTGTGATAGTATTGAAAGAGGTTCTGCTG	1630	
QY	576	GIUcIuLeuGlnValPheLysValProThrProLileThrGlnMetGlnIleSerValLys	595	
Db	1631	GAAAGAAATGACAGTTTTCGGGAACGACAGCTGATTTTCAACATGAGCACTTTTCCATAG	1690	
QY	596	ArgGlnMetLeuTyValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnLys	*615	
Db	1691	CAGCAACAACATATATATGTGTCAACGGCGGGGTGGCCACAGCTCCCTTACACGGGTGT	1750	
QY	616	GluThrTyrlGlyThrIaCysAlaGlnCysCysLeuAlaAlaArgAspProTyGlyAlaTrp	635	
Db	1751	GATATTTAGCGGAAGCGGTGTCTGTAGTGTCTCTCCGCGCGAGACCCCTTACTGTGGTGG	1810	
QY	636	AspGlyLysSerTyThrHisTyArgProSerLeuGlyLysArgArgPheArgArgGln	655	
Db	1811	GATGGTCTGCAATGTTCTGTGCTATTTTCCCACT---GCCAAAGACGCACAAAGACGACA	1867	
QY	656	AspIleArgHisGlnLysProAlaLeuGlnLysLeu-----	667	
Db	1868	GATATATAAGAAATGGAAGACCCAGTCACTGCTGAGATTACACATGATAAATACACAT	1927	
QY	668	GlyGlnSerGlnGlnGlnGlnAlaValGlyLeuValAlaAlaIleThrMetValTyrlGlyThr	687	
Db	1928	GCCCAACGCCCTTAAGAAGAGA-----ATATGATATGGGTGTA	1963	

[illegible]


```
Db 247 TTTTCATTCGACCTGGTAAATATC---AAGCATTTTCAAAAGATGTGTGGCCAGATATCT 303
OY 192 ProGlyGlnArgGluGluCysValArgLysGlyArgAspProLeuThrGluCysAlaAsn 211
Db 304 TACACCAAAAGAGATGATGCAAGTGGCGTGAAGAACATCTGAAAGAAATGTCTAAT 363
OY 212 PheValAlaGluValLeuGlnProHisAsnArgThrHisLeuLeuAlaCysGlyThrGlyAla 231
Db 364 TTCATCAAGGTAAGTAAAGCATATATACAGATCCTGTACCGCTGTGGAGCGGGGCT 423
OY 222 PheGlnProThrHisCysAlaLeuThrValGlyHisArgGlyGlu---HisValLeuHis 250
Db 424 TTTTCATCCAAATTTGCACCTACATTAATAATGGACATCATCCCGAGACAAATTTTAAAG 483
OY 251 LeuGluProGlySerValGluSerGlyArgGlyArgCysProHisGluProSerArgPro 270
Db 484 CTGGAGAACTCACAATTTTGAAGAGCGCGTGGGAAGAGTCCATATGACCCCTAAGCTGTG 543
OY 271 PheAlaSerThrPheIleAspGlyGluLeuGlyThrGlyLeuThrAlaAspPheLeuGly 290
Db 544 ACAGCATACCTTTAATATGATGAGATATATCTGTGAACTGCACCTGATTTATGGGG 603
OY 291 ArgGluAlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAspSer--- 309
Db 604 CGAGACTTGTGATCTCTCGCACTGTGGCAACACCCCAATCAGACAGACAGCAT 663
OY 310 AspGlnSerLeuLeuHisAspProArgPheValMetAlaAlaArgIleProGluAsnSer 329
Db 664 GATTCAGAGTGCCTCAATATGATCCAAAGTTCATATAGTCCACCTCATCTCAGAGATGAC 723
OY 330 AspGlnAspAsnAspLysValTyrPhePhePheSerGluThrValProSerProAspGly 349
Db 724 AATCCTGAAGATGACAAAGTATATCTTTCTCTCCGTGAAT-----GCAATAGATGGA 777
OY 350 ---GlySerAsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGly 386
Db 778 GAACACTCTGGAAGAAAGCTACTACGCTAGAAATAGTGCATATGCAAGAAATGACCTTTGGA 837
OY 369 GlyGlnArgValLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSer 388
Db 838 GGGCAGACAGAGTGGTGAATATGATGACAAACATCTCTCAAGAGCTCTCGATTTGCTCA 897
OY 389 ValProGlyProGlyGluAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeu 408
Db 898 GTGCCAGTCCAAATGTCATGACATCTTTGATGATGATGATGATGATGATGATGATGATG 957
OY 409 TrpProGlyAlaGlyLysSerLeuGluValTyrAlaLeuPheSerThrValSerAlaVal 428
Db 958 AACTTTAAAGATCCTAAATAATCCAGTGTATATGAGCTGTGTACGACTTCCAGTAACATT 1017
OY 429 PheGlnGlyPheAlaValCysValTyrHisMetAlaAspIleTrpGluValPheAsnGly 448
Db 1018 TTCAAGGGATACGCCGTGTATATATGATGATGATGATGATGATGATGATGATGATGATG 1077
OY 449 ProPheAlaHisArgAspGlyProGlnHisGlnTrpGlyProGlyGlyLysValPro 468
Db 1078 CCATATGCCCAAGAGATGAGCCCAATATGATGATGATGATGATGATGATGATGATGATG 1137
OY 469 PheProArgProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheGly 488
Db 1138 TATCCAGGGCCAGAGACTGTGCCAGCAAAACATTGGT-----GCTTTTGAC 1185
OY 489 SerThrLysAspTyrProAspGlyValLeuGlnPheAlaArgAlaHisProLeuMetPhe 508
Db 1186 TCTCAAAAGACCTTCCGATGATGATTAACCTTTGCAAGAAGCATCCAGCCATGTAC 1245
OY 509 TrpProValArgProArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnGln 528
Db 1246 AATCCAGTGTCTCTATATCAATATGCCCAATATGATGATGATGATGATGATGATGATG 1305
OY 529 LeuHisGlnIleValAlaAspArgValGluAlaGluAspGlyThrTyrAspValIlePhe 548
```

```
Db 1306 TTTACCAAAATTTGCTAGACCGAGTGGATGCAAGATGACAGTATGATGATTATGTTT 1365
OY 549 LeuGlyThrAspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlyLysAla 568
Db 1366 ATCGAAGACAGATGTTGGACCCCTTCTTAAGATGATTTCAATTCCTTAAGAGACTGGTAT 1425
OY 569 GluProGluValValLeuGluGluLeuGluValPheLysValProThrProIleThr 588
Db 1426 GATTTAGAAAGAGTCTCTGTCGAGAAATGACAGTTTTTCGGGAACCGACTGATTTTA 1485
OY 589 GluMetGluIleSerValLysArgGlnMetLeuThrValGlySerArgLeuGlyValAla 608
Db 1486 GCATATGACCTTCCATAGCAGCAACATATATATGTTGTTACACCGCTGGGGTGGC 1545
OY 609 GlnLeuArgLeuHisGlnCysGluThrTyrGlyThrAlaCysAlaGluCysCysLeuAla 628
Db 1546 CAGCTCCCTTTACACCGGTGTATATTACGGGAAGCGGTGCTGATGTTGCTCGCC 1605
OY 629 ArgAspProTyrCysAlaTrpAspGlyAlaSerCysThrHisTyrArgProSerLeuGly 648
Db 1606 CGAGACCTTACTGTCTGTGGATGCTGTGCATGTTCTCGTATTTTCCCACT---GCA 1662
OY 649 LysArgArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeu--- 667
Db 1663 AAGAGACGCAAGACAGACAGATATACAAATGAGACCCACTGACTGACTGCTTCAGAC 1722
OY 668 -----GlyGlnSerGlnGluGluAlaValGlyLeuValAla 680
Db 1723 TTTACCATGATATATACCATGACCATGCGCACACGCTGAGAGAGA----- 1764
OY 681 AlaThrMetValTyrGlyThrGluHisAsnSerThrPheLeuGluCysLeuProLysSer 700
Db 1765 -----ATCATCTATGCTGTAGAGAAATAGTACCATTTTGGAAATGACGTCCGAACTG 1818
OY 701 ProGlnAlaAlaValArgTyrPheLeuGlnArgProGlyAspGluGlyProAspGlnVal 720
Db 1819 CAGAGACCGCTGCTCTATTGGCAATTCACAGAGCGGAATGAAAGCCAAAGAAAGATC 1878
OY 721 LysThrAspLysValLeuHisThrGlnArgGlyLeuLeuPheArgGlyLeuSerArg 740
Db 1879 AGAGTGGATGATCATATCATCAGCAGACAGATCAAGCCCTTCTGTACTGATGATCAACAG 1938
OY 741 PheAspAlaGlyThrTyrThrCysThrThrLeuGluHisGlyPheSerGlnThrValVal 760
Db 1939 AAGGATTCAGGCAATTACCTCTCCATGCGGTGGAACATGGTTCATACAAACCTCTT 1998
OY 761 ArgLeuAlaLeuValIleValAlaSerGlnLeuAspAsnLeuPheProGluPro 780
Db 1999 AAGGTAAACCTCGAAGCATATGACACAGACATTTGGAAAGAACTTTCATTAAGATGAT 2058
OY 781 LysProGluGluProProAlaArgGlyGlyLeuAlaSerThrProPro-----LysAla 798
Db 2059 GATGAGATGCTCTAAGACCAAGAAATGTCATATGACATGACACCTTACGCAAGGTC 2118
OY 799 TrpTyrLysAspIleLeuGlnLeuIleGlyPheAlaAsnLeuProArgValAspGlyTyr 818
Db 2119 TGTATACAGAGCTTCTGATCATGATCAACCAACCATTCACACAGATGATGATGATG 2178
OY 819 CysGluArgValTyrCysArgGlyThrThrGluCysSerGlyCysPheArgSerArgSer 838
Db 2179 TGTGAACAAGTTTGGAAAGAGGACCGCAAAACA-----GCTGGGCAA 2220
OY 839 ArgGlyLysGlnAlaArgGlyLysSer-----TrpAlaGlyLeuGluLeuGlyLysLys 856
Db 2221 AGCGCAGACATACCCCGAGGACAGTAAACAAATGAGAACACTTTCAGAAATAATAGAAA 2280
OY 857 MetLysSerArgValHisAlaGluHisAsnArgThrProArgGluVal 872
Db 2281 GSTAGAAACAGAGAGACCCACAGAAATTTGAGAGGGACCCAGAGATGTC 2328
RESULT 13
AAx89112 standard; DNA; 2331 BP
```

XX AAX89112;
 AC
 XX 14-SEP-1999 (first entry)
 DT
 XX
 DE Human brain tissue-derived polypeptide coding sequence (clone OM007).
 XX
 XX Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
 KM recombinant; diagnosis; treatment; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09933873-A1.
 XX
 PD 08-JUL-1999.
 XX
 PF 25-DEC-1998; 98MO-JP05952.
 XX
 PR 26-DEC-1997; 97JP-0358811.
 XX
 PA (ONOV) ONO PHARM CO LTD.
 XX
 PI Fukushima D, Shibayama S, Tada H;
 XX
 DR WPI: 1999-419088/35.
 XX
 DR P-PSDB; AAY27127.
 XX
 PT New adult human brain tissue-produced polypeptides useful for
 PT diagnosis and treatment
 PS
 PS
 PS
 XX
 XX
 CC The invention provides polypeptides (AAY27127-Y27133) produced by human
 CC adult brain tissue, human bone marrow or a human umbilical cord venous
 CC endothelial cell. Host cells transformed with vectors comprising the
 CC nucleic acids encoding the polypeptides are used for the recombinant
 CC expression of the polypeptides. The polypeptides can be used in
 CC diagnosis, treatment and basic studies. With wide applications in
 CC treatment depending on the activity to be aimed at. Sequences
 CC AAX89112-125 represent nucleic acids encoding the polypeptides.
 XX
 SQ Sequence 2331 BP; 723 A; 476 C; 529 G; 603 T; 0 other;
 Alignment Scores:
 Pred. No.: 7 77e-86 Length: 2331
 Score: 1721.00 Matches: 346
 Percent Similarity: 63.89% Conservative: 144
 Best Local Similarity: 45.11% Mismatches: 229
 Query Match: 36.26% Indels: 48
 DB: 20 Gaps: 17
 US-09-813-290-2 (1-875) x AAX89112 (1-2331)
 QY 122 SerValProArgLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArgSerAlaIle 141
 DB 121 AATATTCGAAGACTCAAGCTAACCTCAAAAGACTCTGCTTCAATAGCTATATCC 180
 QY 142 PheLeuArgProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArgAsp 161
 DB 181 TTTTGGGTTTCACAGAGAGCTGGATTTTCAAACTCTCTCTTATGATGAGAAAGAGCC 240
 QY 162 ArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaIleTyrPro 181
 DB 241 AGGCTGCTTGGAGGCCAAGACCATCTTCTACAGTCTGAGTGTACTTAAACAA 300
 QY 182 AspProArgGluValLeuTyrProGlnProGlnArgGlnArgGlnGlnGlnGlnGlnGln 201
 DB 301 AATTTTAAGAAATTTATTTGCTCTGCAAGAAAGAGGCTGAATTTATGTAATAGCT 360
 QY 202 GlyArgAspProLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 221
 DB 361 GGGAAAGATGCCAATACAGAAATTCGCAATTCATCAGAGTACTTACGCCCTATACAA 420

QY 222 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuIleThrVal 241
 DB 421 ACTCACAATATATGTGTGGAACTGGAGCATTTTCATCCAAATATGGGATATATGATCTT 480
 QY 242 Gly---HisArgGlyGlnHisValLeuHisLeuGlnProGlySerValGlnSerGlyArg 260
 DB 481 GGAGTCTCAAGAGGAGATTTATTTCAAACTAGACACAGCTAAATTTGGAGTCTGGCAGA 540
 QY 261 GlyArgCysProHisGlnProSerArgProPheAlaSerThrPheIleAspGlyGlnLeu 280
 DB 541 CTGAATGCTCTTGCATGCTGAGAGCTTTTGTCTTGAATGATGACAGATGAGTACCTC 600
 QY 281 TyrThrGlyLeuThrAlaAspPheLeuGlyArgGlnAlaMetIlePheArgSerGlyGly 300
 DB 601 TACTCTGGACAGCTTCTGATTTCTTGGCAAGATCTGATTCATCAGTACCTGCTGG 660
 QY 301 Pro-----ArgProAlaLeuArgSerAsp---SerAspGlnSerLeuLeuHisAsp 316
 DB 661 CCTACTCATGACCCACCATACATGACAGACTATTTGAGACACTAGCTGCTCAATGA 720
 QY 317 ProArgPheValMetAlaAlaArgIleProGlnAsnSerAspGlnAspAsnAspVal 336
 DB 721 GCAAAATTTATTTGGAACCTTCTTCATACACAGACACCTTCATCCAGATGATGAATA 780
 QY 337 TyrPhePheSerGlnThrValProSerProAspGlyGlySerAsnHisValThrVal 356
 DB 781 TATTTCTTCTTCTGCTGATCA-----TCTCAAGAGAGCACTACCTCGATMAAACATC 834
 QY 357 ---SerArgValGlyArgValCysValAsnAspAlaGlyGlnArgValLeuValAsn 375
 DB 835 CTTCCTCGAGTTGGAGAGGTTTGAAGATGATGTAGAGAGACACACAGCTGATTAAC 894
 QY 376 LysThrSerThrPheLeuValArgLeuValCysSerValProGlyProGlyGlyAla 395
 DB 895 AAGTGGACGATTTTCTTAAAGCCAGACTGATTTCTCAATCTCTGGAAGTGAAGGCGCA 954
 QY 396 GluThrHisPheAspGlnLeuGlnAspValPheLeuLeuTyrProLysAlaGlyLysSer 415
 DB 955 GATTACTTACTTGTATGAGCTTCAAGATATTTATTTACTCCACACAAAGATGAAGAAT 1014
 QY 416 LeuGlnValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
 DB 1015 CCGTACTATATGAGAGCTTACTACACACAGCTCTTCAAAAGGCTGCTGTGT 1074
 QY 436 ValTyrHisMetAlaAspIleTyrGlnValPheAsnGlyProPheAlaHisArgAspGly 455
 DB 1075 GTGATATGACATGGCTGACATCAGACAGCTTTTATATGCTCATATATGCTCATAGGAAGT 1134
 QY 456 ProGlnHisGlnTyrProTyrGlyGlyLysValProPheProArgProGlyValCys 475
 DB 1135 GCAGACCATGCTGGGAGTATGATGATGAGAGAAATTCCTATTCACAGGCTGTACATGT 1194
 QY 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAsp 495
 DB 1195 CCAAGCAAA---ACCTATGACCA-----CTGATTAAAGCCACCGAGATTTTCCAGAT 1245
 QY 496 GluValLeuGlnPheAlaArgAlaHisProLeuMetPheThrProValArgProGlnHis 515
 DB 1246 GATGTATGATGATTTCAAAAGCGGCACTGTGTATATATAGTCCTGATATCCAGTTGCA 1305
 QY 516 GlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleValAsp 535
 DB 1306 GGGAGCAACAGCTTCAAGAAATCAATGATGATTTACAGATGACACAGATAGTGGAT 1365
 QY 536 ArgValGlnAlaGlnAspGlyThrTyrAspValIlePheLeuGlyThrAspSerGlySer 555
 DB 1366 CATGTATTTGCAAGATATGCGCAGTACGATGATATGTTCTTGGAGACACATGGAAGCT 1425
 QY 556 ValLeuLysValIleAlaLeuGlnAlaGlyGlnSerAlaGlnPro-----GlnGlu 572
 DB 1426 GTCCTCAAAAGTGTGACAGAT-----TCAAGAGAAAGATGATGGAAGG 1473
 QY 573 ValValLeuGlnGlnLeuGlnValPheLysValProThrProIleThrGlnMetGlnIle 592


```

Db 1474 GTAGTCTGAGAGGATTCAGATATTCAGACACTATCATCTTGAACATGGAATTC 1533
593 SerVallyArgGlnMetLeuTyrValIglySerArgLeuGlyValAlaGlnLeuArgLeu 612
1534 TCTCTGAGAGGACACAAATTTGATGCTGGTCCGAGATGATTAAGTTCAGCTCTCTTG 1593
613 HisGlnCysGlnThrTyrGlyThrAlaCysAlaGlnCysCysLeuAlaArgAspProTyr 632
1594 CACAGATGCCACACTTATGGGAAGCTTGGCAGACCTGTGTCTTCCAGACAGACCCCTAC 1653
633 CysAlaThrAspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArgArgPhe 652
1654 TGTCTCTGGGATGGAATGCAATGCTCTCGATATGCTCTACT--TCTAAAGAGAGCT 1710
653 ArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCys----- 666
1711 AGAGCCCAAGATGTAAATATGGGACCCATCACCAGTCTGGGACATCGAAGACAGC 1770
667 LeuGlyGlnSerGlnGlnGlnAlaValGlyLeuValAlaAlaThrMetValTyrGly 686
1771 ATTACTCATGAACCTGCTGATGAAAGGTG-----ATTTTGGC 1809
687 ThrGlnHisAsnSerThrPheLeuGlnCysLeuProLysSerProGlnAlaAlaValArg 706
1810 ATTGAATTTAACTCAACCTTCTGGAAATGTATACCTAAATCCCAAGACAACTATTAA 1869
707 TyrLeuLeuGlnArgProGlyValAspGlyProAspGlnValLysThrAspGlnArgVal 726
1870 TGGTATATCCAGAGGTGACGGGATGACATCGAGAGGATTCAGCCGATGAAGAATC 1929
727 LeuHisThrGlnArgGlyLeuLeuPheArgArgLeuSerArgPheAspAlaGlyThrTyr 746
1930 ATCAAAACGGAATATGGGCTACTGATTCGAACTTGCAGAAAGAGATCTGGATGAT 1989
747 ThrCysThrThrLeuGlnHisGlyPheSerGlnThrValAlaArgLeuAlaLeuValVal 766
1990 TACTGAAAGCCAGAGGACACTTTCATCCACACATGATGAGCTGACTTGAATGTC 2049
767 IleValAlaSerGlnLeuAsnAspAsnLeuPheProProGlnProLysProGlnGlnProPro 786
2050 ATTGAAGATTAACAGATGCAAAAT-----ACCCAGAGGCGAGAGATGAGGGGCGAG 2103
787 AlaArgGlyLysLeuAlaSerThrProProLysAlaTyrTyrLysAspIleLeuGlnLeu 806
2104 GTCAGAGATCTATGGCTGAGTCA-----CGCTGAGATACAAAGACTATCAATCAATC 2157
807 IleGlyPheAlaAsnLeuProArgValAlaAspGlnTyrCysGlnArgValITPCysArgGly 826
2158 CTTAGCAGCCCAAACTTC--AGCCTGACACAGTACGCGAACAGATGTG----- 2205
827 ThrThrGlnCysSerGlyCysPheArgSerArgSerArgGlyLysGlnAlaArgGlyLys 846
2206 -----CACAGGAGAGAGCGGAGAGACGAGAAACAGAGGGGGGCCA 2244
847 SerTyrAlaGlyLeu---GlnLeuGlyLysLysMetLysSerArgValHisAlaGlnHis 865
2245 AAGTGAAGCAGACATGCAAGATGAAGAAAGAAATCGAAGACATCAGACAGACCTG 2304
866 AsnArgThrProArgGlnVal 872
2305 GATGAGCTCCCTAGAGCTGTA 2325

```

RESULT 14
AAA37109 standard: cDNA, 3871 BP.

AC AAA37109;
XX
XX 08-AUG-2000 (first entry)
XX
XX Human PRO1491 (UNC760) cDNA sequence SEQ ID NO:309.
XX

```

KW Human: PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
KW ss.
XX Homo sapiens.
OS WO200012708-A2.
PN 09-MAR-2000.
XX 01-SEP-1999; 99WO-US20111.
PF 01-SEP-1999;
XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 09-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100386.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100712.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 23-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.

```



```

Db 941 CTTTCTCGAGTTGGAGAGTTTGAAGATGATGAGAGACACGACCGTGTAAAC 1000
Oy 376 LysTSPserThrpheLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyAla 395
Db 1001 AAGTGGACGACTTTCTTAAGGCCAGACTGATTGTCATTTCTGGAGAGTATGGGCA 1060
Oy 396 GluThrsIshpAspGlnLeuGluAspValPheLeuLeuProGlySAlaGlyLysSer 475
Db 1061 GATTCCTACTTGTGAGCTTCAAGATATTATTACTCCCAACAGAGATGAAGAAAT 1120
Oy 416 LeuGluValTyralaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
Db 1121 CCGTACTAATATGAGCTTCTTACTACCAACGACTTCATCTTCAAGGCTGTGCTTTGT 1180
Oy 436 ValTyrsIshpAlaAspIleTrpGluValPheAsnGlyProPheAlaIshpArgAspGly 455
Db 1181 GTGTATAGCAGGCTGATCAGACAGCACTTTTAAATGCTCATATGCTCATATGAAGAAAGT 1240
Oy 456 ProGlnIshpGlnGlyProGlyGlyLysValProPheProAlaGlyValCys 475
Db 1241 GCAGACCACTGCTGGGCTGAGATATGATGAGAAATTCCTTATCCAGCGCTGTATCATGT 1300
Oy 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyProAsp 495
Db 1301 CCAAGCAAA--ACCTATGACCCA-----CTGATTAAAGTCCACCGAGATTTCACAGAT 1361
Oy 496 GluValLeuGlnPheAlaArgAlaIshpProLeuMetPheTrpProValArgProAlaGhis 515
Db 1362 GATGTCATCAGTTTCATATAGGGGCACTCTGTATTAAGTCCGTATACCCAGTTGCA 1411
Oy 516 GlyArgProValLeuValLysThrIshpLeuAlaGlnGlnLeuIshpIleValValAsp 535
Db 1412 GGAGGACCAACGCTTCAGAGCAATGATGATTCACAGCTGACACAGATAGTGGGAT 1471
Oy 536 ArgValGluAlaGluAspGlyThrTyraSerValIlePheLeuGlyThrAspSerGlySer 555
Db 1472 CATGTCATTCGACAGAGATGCGCAGTACGATTAATGTTCTTGACAGACATTTGGAACT 1531
Oy 556 ValLeuLysValIleAlaLeuGlnAlaGlySerAlaGluPro-----GluGlu 572
Db 1532 GTCCTCAAGTTGTCAGCAT-----TCAAGAGAAAGTGAATATGAGAGAG 1579
Oy 573 ValValLeuGlnGluLeuGlnValPheLysValProThrProIleThrGlnMetGluIle 592
Db 1580 GATAGCTCTGAGGAGTTCGAGATTTCAAGCACTCATCATCTTGAACATGGAATGG 1639
Oy 593 SerValLysArgIshpMetLeuTyraGlySerArgLeuGlyAlaAlaGlnLeuValGlu 612
Db 1640 TCTCTGAGCAGCAACAATGTGATGTTCCCGAGATGATTAAGTTCACTCTCTG 1699
Oy 613 HisGlnCysGluThrTyraGlyThrAlaCysAlaGluCysCysLeuAlaAspProTyra 632
Db 1700 CACAGATGCGCACCTTATGGAAGCTTGGCAGACTGTTGCTCTGCGCAGACGCCCTAC 1759
Oy 633 CysAlaIshpAspGlyAlaSerCysThrIshpTyraArgProSerLeuGlyLysAlaArgPhe 652
Db 1760 TGTGCTGGAGTGAATGATGCTCTGATATGCTCTACT--TCTAAAGAGAGACT 1816
Oy 653 ArgArgGlnAspIleArgHisGlyAspProAlaLeuGlnCys-----666
Db 1817 AGACGCCAAGATGTAAATATAGGCGACCAATCCACCGTCTGGGACATGCAAGACAGC 1876
Oy 667 LeuGlyGlnSerGlnGluGluAlaValAlaGlyLeuValAlaAlaThrMetValTyraGly 686
Db 1877 ATTATGATCATGAACACTGCGATGAAGAGTG-----ATTTTTGGC 1915
Oy 687 ThrGluHisAsnSerThrPheLeuGlnCysLeuProLysSerProGlnAlaAlaValArg 706
Db 1916 ATTGAATTTAATCAACCTTTCTGGAATGTATACCTAATATCCCAACAGCACTTTTAA 1975
Oy 707 TrpLeuGlnArgProGlyAspGlnGlyProAspGlnValLysThrAspGluValArg 726

```

```

Db 1976 TGTATATCCAGAGGCTGAGGAGATGACATTCGAGAGAGTGAAGCCCGCATGAAGATC 2035
Oy 727 LeuHisThrGluArgGlyLeuLeuPheAlaArgLeuSerArgPheAspAlaGlyThrTyra 746
Db 2036 ATCAAAACGGAAATATGGGCTACTGATTCGAGATTTCAGAGAAAGAGATTCGGGATGAT 2095
Oy 747 ThrCysThrThrLeuGlnHisGlyPheSerGlnThrValValArgLeuAlaLeuValAla 766
Db 2096 TACTGCAAGGCCCGAGAGGACACATTCATCCACACATAGTGAAGCTGATTTGATGTC 2155
Oy 767 IleValAlaSerGlnLeuAspAsnLeuPheProProGluProLysProGluProPro 786
Db 2156 ATTGAGATGAACAGATGGAATAAT-----ACCAAGAGCGCAGAGCATGAGAGGGGAG 2209
Oy 787 AlaArgGlyGlyLeuAlaSerThrProProLysAlaTrpTyraAspIleLeuGlnLeu 806
Db 2210 GTCAGAGTCTATTTGGCTGACTGA-----CGGTGACATACAAAGACTTACATCAATTC 2263
Oy 807 IleGlyPheAlaAsnLeuProArgValAspGluTyraCysGluArgValTrpCysArgGly 826
Db 2264 CTTACAGCCCAACTTC--AGCCTGACACAGTACTGCGAACAAGATGTGG-----2311
Oy 827 ThrThrGluCysSerGlyCysPheArgSerArgSerArgGlyLysGlnAlaArgGlyLys 846
Db 2312 -----CACAGGAGAAAGCGAGACAGAGAAACAAAGGGGGGCCCA 2350
Oy 847 SerTrpAlaGlyLeu---GluLeuGlyLysLysMetLysSerArgValHisAlaGlnHis 865
Db 2351 AAGTGAACACATGACGAGAAATGAGAAACGAAATCGAAGACATCAGAGACCTTG 2410
Oy 866 AsnArgThrProArgGluVal 872
Db 2411 GATGAGCTCTCAGACCTGTA 2431
RESULT 15
ID AAS46098 standard; cDNA; 3871 BP.
XX AAS46098;
XX
XX 18-DEC-2001 (first entry)
DE Human DNA encoding PRO polypeptide sequence #174.
XX
XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.
XX
XX Homo sapiens.
OS
XX WO200168848-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 28-FEB-2001; 2001WO-US06520.
PE
XX
XX 01-MAR-2000; 2000WO-US05601.
PR
XX 02-MAR-2000; 2000WO-US05841.
PR
XX 03-MAR-2000; 2000US-187202P.
PR
XX 06-MAR-2000; 2000US-186568P.
PR
XX 14-MAR-2000; 2000US-189320P.
PR
XX 14-MAR-2000; 2000US-189328P.
PR
XX 15-MAR-2000; 2000WO-US06884.
PR
XX 21-MAR-2000; 2000US-190828P.
PR
XX 21-MAR-2000; 2000US-191007P.
PR
XX 21-MAR-2000; 2000US-191048P.
PR
XX 21-MAR-2000; 2000US-191314P.
PR
XX 28-MAR-2000; 2000US-192655P.
PR
XX 29-MAR-2000; 2000US-193032P.
PR
XX 29-MAR-2000; 2000US-193053P.
PR
XX 30-MAR-2000; 2000WO-US08439.

```

PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 03-MAY-2000; 2000US-199654P.
 PR 17-MAY-2000; 2000US-201516P.
 PR 22-MAY-2000; 2000US-2013705.
 PR 30-MAY-2000; 2000US-2014941.
 PR 02-JUN-2000; 2000US-2015264.
 PR 28-JUL-2000; 2000US-209832P.
 PR 22-AUG-2000; 2000US-2020710.
 PR 24-AUG-2000; 2000US-0644848.
 PR 08-NOV-2000; 2000US-0523328.
 PR 01-DEC-2000; 2000US-0530952.
 PR 20-DEC-2000; 2000US-0532678.
 PR 20-DEC-2000; 2000US-0534956.

(GETH) GENENTECH INC.

XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL,
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 DR WPI: 2001-602746/68.

XX P-PSDB; AAU29197.

PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -

XX Claim 2; Fig 347; 774pp; English.

XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
 CC primers for PRO polypeptides of the invention. The sequences of the
 CC invention can be used to detect the presence of a tumour in a mammal by
 CC comparing the level of expression of a PRO polypeptide in a test sample
 CC of cells from the animal and a control sample of normal cells, whereby a
 CC higher level of expression in the test sample indicates the presence of a
 CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
 CC pigs, goats and rabbits but are preferably human. The polypeptides can be
 CC used to stimulate tumour necrosis factor (TNF) alpha release from human
 CC blood, when contacted with it. A specific polypeptide can be used to
 CC stimulate the proliferation or differentiation of chondrocyte cells. The
 CC PRO proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

XX Alignment Scores:

Pred. No.: 1.33e-85 Length: 3871
 Score: 1721.00 Matches: 346
 Percent Similarity: 63.89 Conservative: 144
 Best Local Similarity: 45.118 Mismatches: 229
 Query Match: 36.268 Indels: 48
 DB: 22 Gaps: 17

US-09-813-290-2 (1-875) x AAS46098 (1-3871)

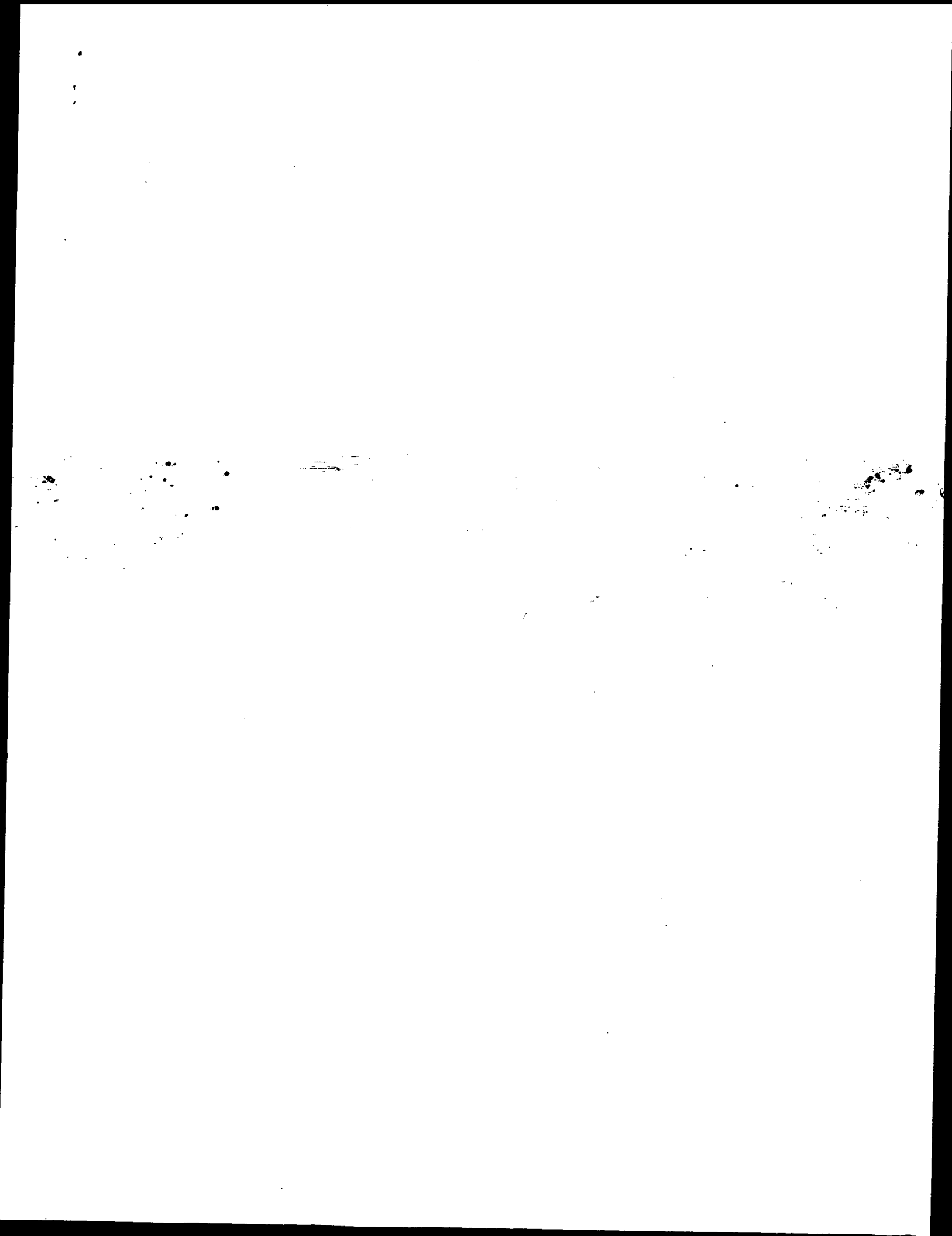
QY 122 SerValProAlaLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArgSerAlaIle 141
 DB 227 AATATTCACAGACTCAAGCTCAAGCTCAAGAGAGCTTCTTCAATAGAGCTGATATTC 286
 QY 142 PheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGlnTyrArgAsp 161

DB 287 TTTTGGTTCATCAGAGAGCTGATTTCAACTCTTCTTAGATGAGGAAGAGGC 346
 QY 162 ArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaIlePro 181
 DB 347 AGGCTGCTCTGGAGGCAAGCAAGACCATCTTCTTACTCAGTGGTGGATTAACAA 406
 QY 182 AspProArgGlnValLeuTyrProGlnProGlyGlnArgGlyGlyValArgGly 201
 DB 407 AATTTTAAAGAAAGTTTATTTGGCTGCTGCAAGAGAGGAGGAGGATTAATTAATTAAGCT 466
 QY 202 GlyArgAspProLeuThrGlyCysAlaAsnPheValArgValLeuGlnProHisAsnArg 221
 DB 467 GGGAAAGATGCCAATACAGATGTGCAAAATTTTCATCAGACTTCCAGCCCTTAACAA 526
 QY 222 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuIleThrVal 241
 DB 527 ACTCACATATATGTGTGTGAGTGGAGCATTTTCATCCATATATATGATGATATCTT 586
 QY 242 Gly---HisArgGlyGlyIleHisValLeuHisLeuGlnProGlySerValGlySerGlyArg 260
 DB 587 GGAGTCTACAGAGGAGATATATATCAACATGACACATATATTTGGAGTGTGGCACA 646
 QY 261 GlyArgCysProHisGlnProSerArgProPheAlaSerThrPheIleAspGlyIleLeu 280
 DB 647 CTGAATGTCTTTTGCATCTCAGACGCTTTTGTTCAGTATGACATGACATGAGTACCTC 706
 QY 281 TyrThrGlyLeuThrAlaAspPheLeuGlyArgGlnAlaMetIlePheArgSerGlyGly 300
 DB 707 TACTGTGGAACAGCTTGTGATTTCTTGGCAAAAGATACGATTCATCTGATCCCTTGGG 766
 QY 301 Pro-----ArgProAlaLeuArgSerAsp---SerAspGlnSerLeuLeuHisAsp 316
 DB 767 CCTACTCATGACACACACTACATACAGAACTGACATTCAGACACTGCTGCTCAATGGA 826
 QY 317 ProArgPheValMetAlaAlaArgIleProGlnAsnSerAspGlnAspAsnArgVal 336
 DB 827 GCMAAATTTATTTGGAACCTTTCTTACATACAGACACTTCAATCCAGATGATGATTAATA 886
 QY 337 TyrPhePhePheSerGlnThrValProSerProAspGlyGlySerAsnHisValThrVal 356
 DB 887 TATTTCTTCTTTCGGAATCA-----TCTCAGAGAGCAGTACTCCGATTAACCAATC 940
 QY 357 ---SerArgValGlyArgValCysValAsnAspAlaGlyGlyGlnArgValLeuValAsn 375
 DB 941 CTTCCTCAGATTTGGAAGCTTTTGAAGATGATGATGAGAGACAGACGCTGTGTAAC 1000
 QY 376 LysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyVal 395
 DB 1001 AAGTGGAGACTTTTCTTAAGCCAGACTGATTTGCTCAATTCCTGGAAGTGAAGGGGCA 1060
 QY 396 GlnThrHisPheAspGlnLeuGlnAspValPheLeuLeuTyrProGlyAlaGlyLysSer 415
 DB 1061 GATACTTACTTGTGATGACCTTCAAGATATTTATTTACTCCCAACAGAGAGTAAAGAAAT 1120
 QY 416 LeuGlnValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCys 435
 DB 1121 CCTGTATATATGAGCTTCTTACTACCAACAGCTCAATCTTCAAGGCTGCGCTTGT 1180
 QY 436 ValTyrHisMetAlaAspIleTrpGlnValPheAsnGlyProPheAlaHisArgAspGly 455
 DB 1181 GTGTATGACATGCTGATCAGACAGAGCTTTTATATGCTCATATGCTCATAGGAAGT 1240
 QY 456 ProGlnHisGlnTyrGlyProTyrGlyGlyValProPheProArgProGlyValCys 475
 DB 1241 GCAGACCATGTTGGGCGAGTATGATGGAGAAATTCCTTATACAGCGCTGTACTGT 1300
 QY 476 ProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAsp 495
 DB 1301 CCAAGCAAA---ACCTATGACCA-----CTGATTAAGTCCACCGAGATTTTCCAGAT 1351
 QY 496 GlnValLeuGlnPheAlaArgAlaHisProLeuMetPheThrProValArgProArgHis 515

Db 1352 GATGTCATCATTTCAATAAGCGCACTGTGATGATATAGTCCGTATACCCAGTTGG 1411
OY 516 G1YArgProValLeuValIysThrHisLeuAlaGlnGlnLeuHisGlnIleValIasp 635
1412 GGAGGACCAACGTTCAAGAGATCAATCATGTGATTACAGTACACAGATAGTGAT 1471
OY 536 ArgValGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 555
1472 CATGTCATTCAGAGATGCGCAGTACGATGATGATGATGATGATGATGATGATGAT 1531
OY 556 ValLeuValIleAlaLeuGlnAlaGlyIleSerAlaGluPro-----GluGlu 572
1532 GTCCTCAAGCTTCACACAT-----TCAAGAGAAAGTGAATATGGAAG.1579
OY 573 ValValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 592
1580 GATGTCGTGGAGAGTTCAGATTCACAGCAGTCATCATCATCATCATCATCATCATCAT 1639
OY 593 SerValIysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeu 612
1640 TCTCTGAAGCAGCAACATTCATGATGATGATGATGATGATGATGATGATGATGATG 1699
OY 613 HisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 632
1700 CACAGATCCGACACTTATGAGAAAGCTTGGCAGACTGTTGTTGTTGTTGTTGTTGTTG 1759
OY 633 CysAlaTyrPaspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArgArgPhe 652
1760 TGTGCTGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1816
OY 653 ArgArgGlnAspIleArgHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 666
1817 AGACCCCAAGATTAATAATATGCGACCAATCCACAGTCTGGGACATCGAAGACAGC 1876
OY 667 LeuGlnGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 686
1877 ATTGATCATGAACGTCGATGAAGAGTG-----ATTTTGGC 1915
OY 687 ThrGlnHisAsnSerThrPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 706
1916 ATTGAATTAACTCAACTTCTGCAATGATATACCTAAATCCACACACACACACTATTAA 1975
OY 707 TyrPheLeuGlnArgProGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 726
1976 TGGTATATCCAGAGTGCAGGAGTGCAGTGCAGAGAGTGCAGAGAGTGCAGAGAGTGC 2035
OY 727 LeuHisThrGlnArgGlyLeuLeuPheArgArgLeuSerArgPheAspAlaGlyThrTyr 746
2036 ATCAAAACGGAATATGGCTACTGATTCGAACTTGCAGAAAGAGATTCCTGGAGTAT 2095
OY 747 ThrCysThrThrLeuGlnHisGlnIlePheSerGlnThrValAlaArgLeuAlaLeuVal 766
2096 TACTGCAGAAACCCAGAGAGCAGCTTTCATCCACACCATGAGTGAAGTGAAGTGAAGT 2155
OY 767 IleValAlaSerGlnLeuAspAsnLeuPheProGlnIleProGlnIleProGlnIlePro 786
2156 ATTGAGAAATGACAGATGAATAAT-----ACCAGAGGCGAGAGCATGAGAGGCGCAG 2209
OY 787 AlaArgGlyIleLeuAlaSerThrProGlnIlePheTyrIleValAspIleLeuGlnIle 806
2210 GTCAGAGATCTATGGCTGATCA-----CGGTGAGATTCAGAAAGACTACATCCAAATC 2263
OY 807 IleGlyPheAlaAsnLeuProArgValAspGlnIleCysGlnArgValTyrCysArgGly 826
2264 CTTAGCAGCCCAACTTC---AGCTCGACAGTACGCAAGAGATGCTG-----2311
OY 827 ThrThrGlnCysSerGlyCysPheArgSerArgSerArgGlyLysGlnAlaArgGlyLys 846
2312 -----CACAGGAGAGAGCGAGAGACAGAGAACAGAGGAGGCGCEA.2350
OY 847 SerTyrAlaGlyLeu-----GluLeuGlnIleLysLysMetLysSerArgValHisAlaGlnHis 865
2351 AAGTGAAGCAGATCCAGAAATGAGAAAGAAATCGAAGACATCAGAGACCTG 2410

OY 866 AsnArgThrProArgGlnVal 872
Db 2411 GATGAGCTCCCTAGAGCTGTA 2431

Search completed: October 9, 2002, 20:19:10
Job time : 348 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 14:12:18 ; Search time 44 Seconds
(without alignments)
1910.867 Million cell updates/sec

Title: US-09-813-290-2

Sequence: 1 MACALGKGVFPMSWPMHK.....KMSRVHAENRRPREVEAT 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1862	39.2	772	2 A49069	collapsin - chicken
2	1832	38.6	772	2 I48747	semaphorin D - mou
3	1830	38.6	771	2 D49423	semaphorin III pre
4	1747	36.8	749	2 G01856	semaphorin V - hum
5	1698.5	35.8	666	2 I58169	semaphorin III - m
6	1658	34.9	753	2 G02173	semaphorin III fam
7	1653.5	34.8	748	2 I48744	semaphorin A - mou
8	1568	33.0	751	2 I48748	semaphorin E - mou
9	862	18.2	834	2 S66498	M-sema F protein p
10	831.5	17.5	782	2 I48746	semaphorin C - mou
11	676.5	14.3	760	2 I48745	semaphorin B - mou
12	628	13.2	730	2 JH0798	fasciclin IV preu
13	606	12.8	724	2 C49423	semaphorin I preu
14	601.5	12.7	656	2 B49423	semaphorin I * fru
15	589	12.6	712	2 T27165	hypothetical prote
16	540.5	11.4	711	2 A49423	semaphorin I preu
17	539.5	11.4	1074	2 JCS928	semaphorin F preu
18	441.5	9.3	653	2 T03102	semaphorin homolog
19	321.5	6.8	676	2 JCS3853	hypothetical prote
20	217.5	4.6	1894	2 JCS980	plexin A precursor
21	190	4.0	1945	2 T13937	plexin A - fruit f
22	187.5	4.0	1884	2 JCS975	plexin 2 precursor
23	185.5	3.9	1905	2 T5153	plexin - African c
24	170	3.6	1568	2 T09074	semaphorin recepto
25	162	3.4	403	2 E42521	A39R protein - vac
26	162	3.4	2051	2 T13164	plexin B - fruit f
27	155	3.3	441	2 S29921	hypothetical prote
28	138.5	2.9	1872	2 JCS976	plexin 3 precursor
29	133	2.8	3375	2 T19621	hypothetical prote

30	132.5	2.8	295	2 J01775	Sall9R protein - v
31	117.5	2.5	1033	2 S19247	cell adhesion prot
32	116	2.4	686	2 B75267	polyol endopeptida
33	116	2.4	1198	2 T28678	polyketide synthas
34	115.5	2.4	2082	2 T37056	probable multi-dom
35	113.5	2.4	2647	2 A37098	gelatin factor AB
36	113	2.4	952	2 S32954	hypothetical prote
37	112.5	2.4	620	2 T30765	hypothetical prote
38	111.5	2.3	6642	2 T29757	protein UNC-89 - C
39	111	2.3	678	2 C68495	hypothetical prote
40	111	2.3	678	2 H72128	3-methyl-2-oxobuta
41	111	2.3	1330	2 S49010	embryonic receptor
42	111	2.3	5175	2 T20992	hypothetical prote
43	111	2.3	5198	2 T43290	hemictin precurs
44	107.5	2.3	1276	2 T18526	SREBP cleavage act
45	107.5	2.3	2611	2 T14591	actinomycin synthe

ALIGNMENTS

RESULT 1

A49069
collapsin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
C:Accession: A49069
R:Uno, Y.; Raible, D.; Raper, J.A.
Cell 75, 217-227, 1993

A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neu
A:Reference number: A49069; MUID:94006554
A:Accession: A49069

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-772 <LHO>

A:Cross-References: GB:002528; NID:9410078; PIDN:AAC59638.1; PTD:9410079
C:Superfamily: semaphorin

Query Match

Best local similarity 39.2%; Score 1862; DB 2; Length 772;
Best local similarity 46.8%; Pred. No. 4.6e-135;
Matches 370; Conservative 135; Mismatches 245; Indels 40; Gaps 12;

QY	103	WLG-----GILLHGSSPSPSPVRLRLSLRLLSANSALFLGPGSLNLAQMYL 156	
DB	3	WLGILLSLGVLLAGRVNQHVNKNVPRKLISKLEESNNIVNENGLANSSYHTFLL 62	
QY	157	DEYRDLFLGLALALYSLRDQAMPREVLPPOGQRECVKGRPLTECANEFVYL 216	
DB	63	DEKRSRLVGAKHHSFNLVNI-KEYQKTVVSHSRDECKWAGKDLRECANFIKVL 121	
QY	217	QPNRTHLACGTGAFQPTCALITVG-HRGEHYLHPGVSVEGRCRHPSPRFASTF 275	
DB	122	KTVNQHLYVACGTGAFHPMCTYLEVGSHPEDNIFRMEDSHFENGKSPYDPKLLASL 181	
QY	276	IDGELVGLTADFLGRAMLFRRSGRRPALRSDS-DQSLHHPFVAARLPENSDDND 334	
DB	182	VDEELVSGTADFGKDFALFTLGHHPRTEDHDSRWMLNDPFIASHLIPESDNEED 241	
QY	335	KYVFESSETPSPDG-GSNHVTSRVGVNDAGGQRLVNNKSTFLKARLVCSVPG 393	
DB	242	KYVFESSETPSPDG-GSNHVTSRVGVNDAGGQRLVNNKSTFLKARLVCSVPG 393	
QY	394	GAETHFDLEDVFLPKRAGKSLVLALESTVSAVPOGFAVCYVHMDIMEVFGPFAHR 453	
DB	300	GIDHFELODFVFLMNSKDRNPITYGVFTSSNIFGSAVCMYSDVRRVFLGPFAHR 359	
QY	454	DGPOHMGPRYGGKVPFPGVCPSPKMTAOPRPGSTKDPDEVLOPFAHPLFWFVR 513	
DB	360	DGPOHMGPRYGGKVPFPGVCPSPKMTAOPRPGSTKDPDEVLOPFAHPLFWFVR 513	
QY	514	RHGPRVLVTKHIAQQLHIVVDRAEDGTVDVIFLIGSSVLKVLALQAGSABEPEV 573	
DB	416	INSPRIMIKTVDVQFTQIVVDRAEDGTVDVIFLIGSSVLKVLALQAGSABEPEV 573	

[illegible]

RESULT 2
I48747
semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48747
R:Puschel, A.W.; Adams, R.H.; Betz, H.
N:Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family, and creates
A:Reference number: I48744; MUID:95267431
A:Accession: I48747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-772 <RSS>
A:Cross-references: EMBL:X85993; NID:g854329; PIDN:CAA59985.1; PID:g854330
C:Genetics:
A:Gene: semD
A:Superfamily: semaphorin

```

Query Match          :   38.6% ; Score 1932; DB 2; Length 772;
Best Local Similarity : 46.4%; Pred No. 9, 3e-133;
Matches    365; Conservative 135; Mismatches 251; Indels      36; Gaps     11;

QY       101 ICWLLGGLLHCGSSGSPFSPVPRILRSLRYDLLSANRSALFIPOGSINTLOAMYDEXR 160
              |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dg        7 IACLPMGVLLTARAYANKNNVPRLKLSYEKLSEMNNTFTNGLANSSSYHTFLDEER  66
               |||||
QY       161 DLEFGGDALTASLRDQAMPDPREVIWMPROGORECEYRKRRDLTECANFRVLIOPN  220
              ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db         67 SRLYYAGADHLTFENLVNI-KDFOKIWEVSYSYTRDECKEMAKDKILKECANFKIVLEYA  125
              ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY       221 RTHLLACGTGAFOPTCALITVHGRC-HYLHLEPSVESGRCPHEPGRPASTPEIDEE  279
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        126 QTHLYACSTGANFHPICTYLEVGHPEDNIEFKQDBHFENGKRKSYPDRKLLASLTIDBE  185
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY       280 LYTGLTAFLCREAMIFESGGPRPALRSDS-DQSLLHDPREVAARIPENSDOODNDKYVE  338
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        186 LYSGTADEFMRDAIFETPLDHPIRTFOHDSRWLPNRFISAHILPESDNPEDDKYVE  245
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY       339 FFSETVESPDCG--SNHYTSRVGCVCVDNAGGCVLNKKMSTFFLKARLYCVCBPBGCAE  396
              ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        246 FFE-----NAIGEGSHGKTKHARTIGQICANDRGHGHSNLNKWTFFIKARLICSVPPNGID  302
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY       397 THFDQLEDVILLMERACKSLEEVYALFESTVASVFQGFACVMADMIEWFNCFPAHRGP  456
              ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        303 THFDQLDVLFMMNSKDKNRPILVYGVTFTSSNIPKSAVCAMCMMSBVRVRFELGYAHNRGP  362

```

[illegible]

```

RESULT 3
D49423
semaphorin III precursor - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: D49423
R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
Cell 75, 1389-1399, 1993
A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth co
A:Reference number: A49423; MUID:94094332
A:Accession: D49423
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-771 <KOL>
A:Cross-references: GB:126081; NID:g799328; PIDN:AAA65938.1; PID:g436560
C:Genetics:
A:Gene: GDB:SEMA1
A:Cross-references: GDB:283448
;Superfamily: semaphorin

```

[illegible]

Oy		254	GVSYSGCRPDHESRPFPAFTFIDIGELYLTADFLQREMIKRRSGPRALSD--SDDS	312
Db		159	GRLEDGKGKSPYDRIHRNAASVLGEELVSGVAADLMQRDTIFRSLGQPSRLTEPHDSR	218
Oy		313	LIMDPRYMAARIPENDODNDKYFFFESETVSPSDGGSNNHYVSRVGRGVMDAGGORY	372
Db		219	WLNPKPVKFWPIESENPNDDKIYFFFRETAEADAPALGLRSLSRGGQLCRNDVGORS	278
Oy		373	VYNKSFILKARLVCSPVGPFGAETHFDQLEDFVLLMPKAGKSLEYALPSTVASAFQGF	432
Db		279	LVNKKTFITKLARLYCSVPYGVEG-DTHFDQLDQVFLSSRDHRPLLYAVESTSSLIPOGS	337
Oy		433	AVCYIHMAADIWEVNGPPAHHDGFOHQMGPRGKVFPFRPGVCSKMTAQPRFGSTKD	492
Db		338	AVCVSNMDEVRAFLGPFARKEGPMHWVSQGRVPYPDRDMCPSKTFG---TFS TKD	393
Oy		493	YRDEYLQFARHPFMFVVRPRHGPRVLYVKTHLAQOLHOIYVDVEAEDGTADVILELTD	552
Db		394	FPPDYIQARRHPMLKNVLPFTGSRPLFLQGANVFYFTQIAADRVAADAHDGYDLFIGTD	453
Oy		553	SGSVLYKTAIALOAGSADEEVYVLELOLFKYPRPTITEMEISVKROMLYGSRIGVADRL	612
Db		454	VGTVLKVTISVSKGSRPSAGELLLELHFEDSAAVTSMQISSKHOLYVSRSAVAIAL	513
Oy		613	HOCETYGACAECCICARPYCAMPDCACTHPRPSIGKRFRRODIRHGNPALQCLGSOE	672
Db		514	HRCNAHGKVCIECCILARPYCAMDGACTRPQPS-ARKRFRRQDVNRNDPSTLCGDSR	572
Oy		673	EEAVGLVAATVYVGTENSTFLECIPKSPQAAMRLIORPDEBGPDOYKTDERVLHNERG	732
Db		573	P-----ALLEHHYFEGVESASFLECEPRSLQARVETPTFRAGVTANTOVLAEREKERARG	628
Oy		733	LIFRLSLPFDATGYCTCLEHGFOSOTVRLAVIVASQNLNF-PPEPKDEEPAPAGL	791
Db		629	LILRLRLRRDSSVYLCAAVEGCFOTPLRLSLSHVLSAQABRLARAEEAANAAPG----	684
Oy		792	ASTPRKAWYKDILOLI-----GFANLPVDEYECERWCRCRGTECGSC-FKRSRKGAOR	844
Db		685	----PKIWYRDLQIVFPGGGGSANSLRM-----CRPOPALQSLEPLESRRKGRMR	731
<hr/>				
RESULT 5				
158169 semaphorin III - mouse (fragment)				
C.Species: Mus musculus (house mouse)				
C.Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999				
C.Accession: L58169				
R.Messenger: E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S., et al.				
Neuron 14, 949-959, 1995				
A.Title: Semaphorin III can function as a selective chemoattractant to pattern sensory				
A.Reference number: L58169; MUID:95267432				
A.Accession: L58169				
A.Status: Preliminary; translated from GB/EMBL/DBJ				
A.Molecule type: mRNA				
A.Residues: 1-666 <RES>				
A.Cross-references: GB:L40484; NID:g703189; PID:NAAA73934.1; PTD:g703190				

[illegible]


```

Db 121 ISAHLPESDNEDOKVFFEREN--AIDGESHGKATHARIQOICKNDGGRSLVNMKT 178
QY 379 TELKARLVCSVPGGATHPDQLEDFVLLMPKAKSLEYVALFESTVASVFOGFAVCYH 438
Db 179 TFLKARLVCSVPGGATHPDQLEDFVLLMPKAKSLEYVALFESTVASVFOGFAVCYH 238
QY 439 MADIVEVNGFPAHNDGPOHONGPVGKVPFPPRGVCPSCMAOPGRCFPGSKADPEVL 498
Db 239 MSDVRRVLLGPRYAHNDGPNYQWVQGRVPRYDRPTCSKTTG---GFDSTKDLPDVYI 294
QY 499 QFARHPLMFPPRRRGRPVLLVKTALAQOLHOIVDRAVEDGTYDVIETGDSGSVYLK 558
Db 295 TFASSHPRAMPVPPFNNRPIMIKTDVNYQFQIYVDRVDAEDCOQDYVWFISTDVGTVLK 354
QY 559 VIALOAGSGAPPEEYVLEDFKVPRTITEMEISVKROMLYVSRILVAQOLHOCETY 618
Db 355 VVSPVKEVHMLDEVLLEEMTVFREPITISAMELSTKQOOLYIGSTAGVADQLPHRCOIY 414
QY 619 GTACAECCCLADPYCAMDGASCTHYRPSLGRFRRODIRGNPALQCL-----GQ 669
Db 415 GKACAECCCLADPYCAMDGASCTHYRPSLGRFRRODIRGNPALQCL-----GQ 669
QY 670 SOEEAVGLVATVYVGEHNSTELECPKSPQAAVRLDOPDEGPDQYKTDERVLHT 729
Db 474 SLER-----IYGVNSSTLECSPKSORALVYQFORNDREKEIKMDHIIRT 525
QY 730 ERGLFRRLSFPDAGTYCTLEHGFSTQVRLALVIVASOLDLFPREPKEPREPARG 789
Db 526 EGGLLLSLQKSDGNVNLCHAVENGFMTLKVLEVIDTEHLELLKDDGDSKIKE 585
QY 790 GLASTPP--KAWYDILQIGFANLPRVDEYCEVWRCGTTEGSCGCFRSRSGKQARSG 847
Db 586 MSSSMTPSOKWYVDFDQOLNHPNLNTEWDECEQWKKRQO----RQRGHSGSGSSNK 641
QY 848 WAGLEGGKKKSRVHAHNRRPREV 872
Db 642 WKHMQSKKGNRRTHFERAPRSV 666

RESULT 6
semaphorin III family homolog - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G02173
R:Naylor, S.
submitted to the EMBL data library, October 1995
A:Reference number: G09275
A:Accession: G02173
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-753 <NAY>
A:Cross-references: EMBL:U08276; NID:g1061350; PIDN:AA18276.1; PID:g1061351
C:Superfamily: semaphorin

Query Match 34.9%; Score 1638; DB 2; Length 753;
Best Local Similarity 44.5%; Pred. No. 2.2e-119;
Matches 345; Conservative 127; Mismatches 258; Indels 46; Gaps 18;

QY 104 LIGGLLHGCGSSGPPSG-----PSVRLRLSYRDLLSANRSALITLGOGSLNLOAMY 155
Db 2 LVAGLLIMASLTGAMPSPFTQDHLPRVRLSFKELKATGTANHFNFLLNTTDDYRLL 61
QY 156 LDEYDRLEFLGIDLALYSRLDQAMPDPREVLPPOGQREECVARKGRDPLECANFVRV 215
Db 62 KDBDHRMRYGSKDYVLSLDLHDIINREPLIIHMAASPORIECYLSKDVNGECGNFVRL 121
QY 216 LQHNHTHLLACGTGAFQPTCALITYGHG--EHVHLPEGSVESGRCRPHESRPPAST 274
Db 122 IQPWNHTHLYVCGTGAJNPCTYVNRGRADODYIFYLEPERLESKCKCPYDKPLDTASA 181
QY 275 FIDGELYTGTLADFLGRAMAFIRSGGRPALRSDSQS--LLHDPFVMAARIFENSQDN 333

```

```

Db 182 LINEELIACVYIDFMTDAIFRTLGKQYAMRTDOYNSHMLNDPSTIAELLIPDAEND- 240
QY 334 DKVYEFSETPSPDGGSNHVTSHRGVAVNDAGQVRLVNMKSTFLKARLVCSVPGG 393
Db 241 DKLYEFFERR--SAEAPQSPAYARIGRLCUNDGSGHCLVNMKSTFLKARLVCSVPGG 298
QY 394 CAETHPDQLEDFVLLMPKAKSLEYVALFESTVASVFOGFAVCYHMAIDIVEVNGFPAH 453
Db 299 GIETHFDLQDFVQOQDVRNDVYAVFTSSGVFSGAVCVYSMADIRMFNCFPAHK 358
QY 454 DGPQHQMGVPGKVPFPPRGVCPSCMTAOPGRFSTQDYPREVLOFARHPLMFPPRP 513
Db 359 EGPNNQWMPFSGKMYPRPGTCPG--TTPRS--MSTADYPREVINTMSHPLQVAYP 415
QY 514 RHGRVLYKTLAQLHOIVDRAVEDGTYDVIETGDSGSVLYALQAGSAEPEV 573
Db 416 LQRRPLVVRGTGAPVLTITAVDQVSDAGRYEVLFLGTRGTQVLYLPK--DQGMEL 474
QY 574 VLEELQVKKVPTITEMEISVKROMLYVSGRGLVAQOLHOCETYGTACAECCCLADPYC 633
Db 475 MLEEVEFKDPAVVTWITTSKRQOLYVAVGYTHLHRCQAGAACADCCCLADPYC 534
QY 634 AMDGASCTHYRPSLGRFRRODIRGNPALQCLG--SOEEAVGLVATVYVGEHNS 691
Db 535 AMDGASCTHYRPSLGRFRRODIRGNPALQCLG--SOEEAVGLVATVYVGEHNS 691
QY 692 TFLKARLVCSVPGGATHPDQLEDFVLLMPKAKSLEYVALFESTVASVFOGFAVCYH 750
Db 590 AFLEQCPSPQATYKWLPRQPRGR--RREIRAFEDFLTEGGLLRALQLSDRGYSCYA 648
QY 751 LEHGFSTQVRLALVIVASOLDLFPREPKEPREPARGGLASTPPKAWYDILQIGF 809
Db 649 TENNFKHVYTRVQLVLRDVAHVALEPPLMSAPRPGAG---PPPPYQELAQLLAQ 704
QY 810 ANLPRVDEYCEVWRC--GTTEGSCGCFRSRSGKQARSGWAGLEGGKKKSRVH 862
Db 705 PEVGLIHQYCGYWRHVPSPREAPGARSPEDQO-----KTRRNRRH 748

RESULT 7
semaphorin A - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48744
R:Puschel, A.W., Adams, R.H., Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creat
A:Reference number: I48744; MUID:95267431
A:Accession: I48744
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-748 <RES>
A:Cross-references: EMBL:X85990; NID:g854323; PIDN:CAA59982.1; PID:g854324
C:Genetics:
A:Gene: sema
C:Superfamily: semaphorin

Query Match 34.8%; Score 1653.5; DB 2; Length 748;
Best Local Similarity 46.1%; Pred. No. 4.9e-119;
Matches 358; Conservative 106; Mismatches 250; Indels 63; Gaps 17;

QY 83 GRRCRCPQPPSMASAMALQMLG--GILLHGSSGSPGSPVRLRLSYRDLLSANRSAL 141
Db 2 GRAEAAVMDGL-----ALMVAGLG-----DTAPNLPRLSLSFQ--IQARHGRV 45
QY 142 FLPGOSLNLQAMYLDEYDRLEFLGIDLALYSRLDQAMPDPREVLPPOGQREECVARK 201
Db 46 TFLRERTCGVEALLVDERRLFLVGAENHVASLSDNISRAKRLAMPAYERRECENMA 105
QY 202 GRDPLTCANFVRLQPHNHTHLLACGTGAFQPTCAL---ITVGHREHVLHLEPGSVES 258

```

Db 106 GKDIGTECMNFVRLHVAHNTHTLLACRTGAFHPICALMRMTAG--GTHA-STGPKLELD 162
 QY 259 GGRCPHEPSPPEASTFIDGLYGLTADFLGREAMTRSGGPPALRSD-SDGSLHDP 317
 Db 163 GKGTTPYDPRRPPSVLVEGLYSVGTADLMGRDFTTFRSLGONPSLTETPHDSHMLNEP 222
 QY 318 RFWMAARIPESDNDKNDKVFYFEFSTVPSPDGGSNHYTVSVRGVCVNDAGQAVLVKMK 377
 Db 223 KFAVWVETPESENPDQDITFEFFRESAVEAARPAKGRMSVSRVGOICRNDLGGQRLVVKW 282
 QY 378 STFLKARLVCSVPFGAETHFDLEDFLLMPKAKSLEYALFSTVSAVFGFVAVCY 437
 Db 283 TTFPLKARLVCSVPFGAETHFDLEDFLLMPKAKSLEYALFSTVSAVFGFVAVCY 341
 QY 438 HMADIWEVNGPFAHRDGPQOMGFGYGVKVPFRPGVCPKMTAOPGRFGSTKRYXPEV 497
 Db 342 SMNDVRRAPFLPRLPKHEGPTQHWVSQGRVPRPGMCPKSTEG---TFSTKDPDGV 397
 QY 498 LOFARHPLMPVPRRGRPLVKTHLAOLHQIVVDVREAEDEGTVDVIFLGTSGLV 557
 Db 398 IOFGNHPFLMNPVLPKGRPLFLVGVAGYFTQTADRVAAADGHYDLFICGTVGLV 457
 QY 558 KYIALQAGSAPEEVEVLEELQVFRVPRPTTEMETSVKROMLYVSGRLGVAOLRLHOCET 617
 Db 458 KYISVPRKGRPSGLLLEELQVFEDESAITSMOISSKROOLYVSAARAVALIALHRCFA 517
 QY 618 YGTAAACECLADPCANDGASCTHYRPSLGRRRRRODIRGNALOCLOGSOEEAVG 677
 Db 518 IGRACAECLADPCANDGASCTHYRPSLGRRRRRODIRGNALOCLOGSOEEAVG 572
 QY 678 LVAAGVYSTEHNSTFELCPKSPQAARWMLQVRGDEGPDVKTDERVLTHERGLLFR 737
 Db 573 VLEKLVGVESGSAFLCEPRSLQAHVQWTFQAGAEAHQVLAEEVERIARAGLLLG 632
 QY 738 LSRFDAGTYTCTTLEHGSQIVYRLALVYASQDLNLFPRPKPEPRPARGLASTP-- 795
 Db 633 LRROSGGVLYCAVDEGFSQPLRLRLVHLVLSAQAERL-----ARAEAAAPAP 681
 QY 796 --PKMYKDILOLI-----GFANLPRVDEYCEWRCGTTF-ECSGCFRSRSRKGAR 844
 Db 682 PGPKMYRPLDOLVERGGGGGANSIRM-----CRPGHHSVADSRKGRNR 730

RESULT 8

semaphorin E - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
 C:Accession: I48748
 R:Puschel, A.W.; Adams, R.H.; Betz, H.
 Neuron 14, 941-948, 1995
 A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
 A:Reference number: I48744; MUID:95267431
 A:Accession: I48748
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-751 <RES>
 A:Cross-references: EMBL:X85994; NID:g854331; PIDN:CAA59986.1; PID:g854332
 C:Genetics:
 A:Gene: sem
 C:Superfamily: semaphorin

Query Match 33.0%; Score 1568; DB 2; Length 751;
 Best Local Similarity 42.4%; Pred. No. 1.9e-112;
 Matches 337; Conservative 116; Mismatches 265; Indels 76; Gaps 15;

QY 100 AICMLGLGLLH--GGSSGSPSPSPRLRLSLRDLSSANKSAIFLPGQSLNLMQMYL 156
 Db 5 AICLVGVFICISICVSSQPO---ARVYLTFDELREKTSSEYLSHQOQDYLRLM 59
 QY 157 DEYDRFLGLGLDALYSLRLDQAMPDPREVLMPPQPGQRECVKGRDPLTECANFVRL 216
 Db 60 DEODRIYVSKDHILSLNININISQELPSVFWPASTIKVECKMAGNDPTGCGNFVRI 119

QY 217 QPHNTHLLACGTGAFQPTCALITVGHGEHYLHLEPSSVESGRCRCHPSPASTFI 276
 Db 120 QFNENHLYVCGSGAFSPVCYIYNKGRSSEQGVFMIDKCEGKRCRCSFNPVNTVSMI 179
 QY 277 DGEIYGLTADFLGREAMIFRSGRPPALRSDSDS-LIHDRFWMAARIPENSQDNDK 335
 Db 180 NEELFSCHWIDFMGTDAIFRSLTKRMQLRTOHNSKMLSEMFDAVHIDGTPNDAK 239
 QY 336 YVFESEVPPSPDGSNNHYTVSVRGVCVNDAGGQVRLNKKVSTFLKARLVCSVPFGA 395
 Db 240 YVEFERKRLTDNNRSTKQIH-SMARICPNDDGGORSLNKKTTFLKARLVCSVDEDCP 298
 QY 396 ETHFDLEDFLLMPKAKSLEYALFSTVSAVFGFVAVCYHMADIWEVNGPFAHRD 455
 Db 299 ETHFDLEDFLLMPKAKSLEYALFSTVSAVFGFVAVCYHMADIWEVNGPFAHRD 358
 QY 456 POHQMPYGGKVPFRPGVCPKMTAOPGRFGSTKRYXPEV 515
 Db 359 PNHOLISTYOGRIYPRPGTGPFGAFTPNMR---TTKDFPDVVTFLRNPMTNISTPIH 415
 QY 516 GRPLVKTHLAOLHQIVVDVREAEDEGTVDVIFLGTSGLV 575
 Db 416 RPLVKTHTLAOLHQIVVDVREAEDEGTVDVIFLGTSGLV 474
 QY 576 EELQVFRVPRPTTEMETSVKROMLYVSGRLGVAOLRLHOCETYTACAECLADPCY 635
 Db 475 EELQVFRVPRPTTEMETSVKROMLYVSGRLGVAOLRLHOCETYTACAECLADPCY 534
 QY 636 DGASCTHYRPSLGRRRRRODIRGNALOCLOGSOEEAVGLVADNMYGTEHNSTFE 695
 Db 535 DGASCTHYRPSLGRRRRRODIRGNALOCLOGSOEEAVGLVADNMYGTEHNSTFE 591
 QY 592 CAKSPQASIKMLQDKDKRKEG---KLNERTIATSOGLILRSYODSDQGLHYCIATE 647
 Db 647 CAKSPQASIKMLQDKDKRKEG---KLNERTIATSOGLILRSYODSDQGLHYCIATE 602
 QY 648 NSFKQTIKINFLVLDSEMAVAVTDKWSPTWAGSVRALFPH-----KD 652
 Db 652 NSFKQTIKINFLVLDSEMAVAVTDKWSPTWAGSVRALFPH-----KD 602
 QY 803 ILQILGFANLPRVDEYCEWRCGTTECSGCFRSRSRKGAR 856
 Db 693 ILQILGFANLPRVDEYCEWRCGTTECSGCFRSRSRKGAR 731
 QY 857 MKSRVHAENRTPR 870
 Db 732 LKALINSRKSRRR 745

RESULT 9

S66498
 M-sema F protein precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999
 C:Accession: S66498
 R:Inagaki, S.; Furuyama, T.; Iwahashi, Y.
 FEBS Lett. 370, 269-272, 1995
 A:Title: Identification of a member of mouse semaphorin family.
 A:Reference number: S66498; MUID:95385809
 A:Accession: S66498
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-834 <INA>
 A:Cross-references: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599
 C:Superfamily: semaphorin
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-834/Product: M-sema F protein #status predicted <MAT>

Query Match 18.2%; Score 862; DB 2; Length 834;
 Best Local Similarity 31.6%; Pred. No. 4.8e-36;
 Matches 252; Conservative 106; Mismatches 282; Indels 158; Gaps 30;

QY 94 MARSAMICWLLGGLLHGSSGSPSPSVRLRLSLRDLISANRSLFLGPGSLNQA 153
 DB 1 MAPH-WAV-WLLAAGLWGLGAGEMMMNLVPRKTVSSSELTVVRRFSQTIQDFILITL 57
 QY 154 MYLDEYDRFLFLGDLALYSRLDQ-----AMPDREVLMPPOPGORECVRGRDPL 206
 DB 58 --LTHSGLLVYGAREALFAVSVALEIGAIISWAPAE-----KKICETOKKGSNQ 107
 QY 207 TECANFVRLVLPNHRTHLLACGTAFOPTCALITVGHGHEVHLHLEPGSVESGRCHHE 266
 DB 108 TECNFIRFLQPVNSSLHYVCGTAFQPKCYIIM-----LTFLLDRAEFEDGKCPDYD 162
 QY 267 PSRPFASFTFDGELYTLADFLGREAMIFRSRGPRLRSDOSLHDPREYMAARIP 326
 DB 163 PARGHTGLVLDGELYSLTNLFLCTEPEVILRYMGTHHSIKTEYLAFMLNEPHFVGSFVR 222
 QY 327 ENSDQ---DNDKYFEFSETPSPDGSNNHYTVSRVGRVCVNDAGGGRVLYNKKSTLKA 383
 DB 223 ESVGSTFGDDDKITFEFSERAVEIDCYSEOV-VARVARVCKGDMGARGTLOKKMTTELKA 281
 QY 384 RLVCYVPGGGAETHFDQLEDFVLLMPKAKSLLEVVALFTVSAPFOGFAVCVYHMADIW 443
 DB 282 RLVCAP---DMKYEFNQLKAVHTLRGASWHNNTFEFVQFARMGMDMLSAVCEYQLDQIQ 338
 QY 444 EVNGRPAHRDQOHOMGRPYGKVPFRPGVCSKMTAQGRPGSTKDYDDEVYQARA 503
 DB 339 QVEGPRKEYSEQAKARATYDVPSPRPGSCJNNMHRDNG--YTSLELDPDNTLFIKK 396
 QY 504 HPLMFVPRRHRGRPVLVK-----THLAQOLHOIVDRVNEADG-TYDVFELIGDGSV 557
 DB 397 HPLMEQVQRKRLGRPLLYKNTNFTH-----VADRVGLDGAITYVFLIGDQWL 449
 QY 558 KVALDAGSABEPEVLELOVKFVPTITEMEISVKRMLYVGRVLAVALHOCET 617
 DB 450 KAVSL-----GWIMHVELEQVFD--QEPVESLVLSSQKVLFPAGSSQVLQSLADCTK 502
 QY 618 YGACAECCCLARPDCYCAMDASCYTHRPSLGRK-RFRRODIRRGNPALQCLGSGQSEAV 676
 DB 503 Y-REYDCVLAARPYCAMVNNISRCVATTSGRSGSEFLVQHVANLDTSKMC-----NOYGI 556
 QY 677 GLVAA-----TWYGTGHNSTFLECLPKSPQAAVRML-----ORPG--DEGPDQV 720
 DB 557 KYKRSIPKNTYVSGD--LVLPCHLSNLAHMTFSGQDLPAQPGSLFDTG-----609
 QY 721 KIDERYLHTEGRLFRRLSFDAGTYCTTLENGSOTVRLA-----LVYVA-----769
 DB 610 -----LQALVMAAOSRHSRGPYRCYSEEG-----TRLAESYLAVVAGSSVTLE 655
 QY 770 --SOLDNL-----FPEE-PK-781
 DB 656 ARAPLENLGLVWLAVALAVCLVLLVLSLRRLRELEKAKASERLYVPLELPKE 745
 QY 782 PEEPPARGGLASTPPKAW 799
 DB 716 PASPPFRPG-PETDEKLM 732

RESULT 10

148746

semaphorin C - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 03-Nov-2000

C:Accession: I48746

R:Puschel, A.W.; Adams, R.H.; Betz, H.

A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates

A:Reference number: I48744; MUID:95267431

A:Accession: I48746

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-782 <RES>

A:Cross-references: EMBL:X85992; NID:g854327; PIDN:CAA59984.1; PID:g854328

C:Genetics:

A:Gene: semc
C:Superfamily: semaphorin

Query Match

Best local similarity 32.3%; Pred. No. 9,8e-56;

Matches 223; Conservative 110; Mismatches 281; Indels 77; Gaps 22;

QY 150 NLQAMYLDEYDRFLFLGDLALYSRLDQAM---PDREVLMPPOPGORECVRGRDPL 206
 DB 15 NYTALLLSQDGKTLVYGAREALFALNSNLSFLPGGEYQELMSADADRKKQCCSFKGRDPK 74
 QY 207 TECANFVRLVLPNHRTHLLACGTAFOPTCALITVGHGHEVHLHLEPGSV--ESG 259
 DB 75 RDCQNYIKILLPLNSSLHLLTGTAFAFSLCYI-----HIAFTLAQDPAGVYLEDG 127
 QY 260 RGRCPHEPSRPFASFTFDGELYTLADFLGREAMIFRSRGPRLRSDOSLHDPRE 319
 DB 128 KGHCPFPNFKSTALVVDGELYTGVSFGNDPAISRQSSRPT-KTESLNLMDPAF 186
 QY 320 VMAARFENSQ---DNDKYFEFSETPSPDGSNNHYTVSRVGRVCVNDAGGGRVLYN 376
 DB 187 VASATSESLGSPIGDDDKITFEFSGOEPEFENT-VSRVARVCKGDEGGRVLDQR 245
 QY 377 WSTFLKARLYCSVPGGAETHFDQLEDFVLLMPKA---GKSLVVALFTV--SAVEOG 431
 DB 246 WTSFLKAQOLLCSDPDG---FPFNVLDVFTLNNPDQMKRTLSI-GVFTSQMHKRTTEG 301
 QY 432 FAVCVYHMADIWEYFNRPFAHRDQOHOMGRPYGKVPFRPGVCSKMTAQGRPGSTK 491
 DB 302 SAICVFMTNDVOKAFDGLYKVNRETOQWYETEHQVTPPRACITNSARE--RKINSL 359
 QY 492 DYDEVLQFARAHLPMFWRPRRHRGRPVLVKTHLAQOLHOIVDRVNEADGTYDVFELGT 551
 DB 360 QLPDRVLNFKDHFMLMGQVRSR-----LLLDPRARVQRVAVHRYGHLSTYDVLFTGT 414
 QY 552 DSGSVLYKVALQAGSABEPEVLELOVKFVPTITEMEISVKRMLYVGRVLAVALH 611
 DB 415 GDBRLHRAVTLSS-----RVHIIIEELQIFPGQCPVQNLHDSHGLLVYSSHGAVQVP 468
 QY 612 LHOCETYGTACACCLARDCYCAMDASC---THYRSLCKRFRRODIRRGNPALQCLG 668
 DB 469 VANCSLYPT-CGCDLARDYCAMTGSACRLASLYQDLSARFW-TQDIEGASVKELCXN 526
 QY 669 QSGEEAVGLVAATMYG-----TEHSTFLECLPKSPQAAVRMLQRPDEGPDQV 720
 DB 527 SSYK-----ARFLVPGKCKQVOIQPNTVNTLACLPLSNLAFRLV-----HNGAPVNA 575
 QY 721 KIDERYLHTEGRLFRRLSFDAGTYCTTLENGSOTVRLALVYVASQDLNLEPPEP 780
 DB 576 SASCRVLPF-GDLLVGSQOGLGVQCWSIEBGFQOLVASCYPEVMEEGVMDQKMRDG 633
 QY 781 KP-----EEPPARGGLASTPPKAWKDL 804
 DB 634 TPVIINTSRVASAPAGGRDSMGADKSTWNEFL 664

RESULT 11

148745

semaphorin B - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999

C:Accession: I48745

R:Puschel, A.W.; Adams, R.H.; Betz, H.

A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creat

A:Reference number: I48744; MUID:95267431

A:Accession: I48745

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-760 <RES>

A:Cross-references: EMBL:X85991; NID:g854325; PIDN:CAA59983.1; PID:g854326

C:Genetics:

A:Gene: semB

C:Superfamily: semaphorin

Query Match 12.8%; Score 606; DB 2; Length 724;
 Best Local Similarity 29.0%; Pred. No. 2e-38;
 Matches 187; Conservative 94; Mismatches 260; Indels 104; Gaps 26;

QY 147 GSNLQAMVDEYDRDLFLGLDALLSLRLDQAMPD--PREVL-WPPOGORECVRKGR 203
 DB 64 GKLYYRFHNMEDDTLYGAMDRVFRVNLONISSNCRDAINLEPRDVSVCVSGK 123
 QY 204 DPLETCANFVRLDPHNR-THILACGTGAEPQ-----TCALITVGRGHEVHLHPGSVES 258
 DB 124 SQIFEDCKNHRVLOSOGDRLYVCGTNAHNPQVYIYANLTHLPREXYI-----GVGL 178
 QY 259 GRGRCPHEPRPFASTFIDG-----ELYTGLADFLGREAMIFRSGRPPALRSDS-- 309
 DB 179 GIAKCPDPLDNSTAIYVENGNGPGLGSGTNAEFTKADTVIFRD-----LYNLSAK 233
 QY 310 -----DQSLIHDRFVMAAIPENSDDNDKVFEESET--VPSPDGGSNHYTV 356
 DB 234 RLEKRRRTIKLYSKWLDKRFNEVGSFDGE-----YVYFFRETAVEYINCGR--AVY 284
 QY 357 SRVRCVNDVAGGQRYLVNKMSTFLKARLYCVPGGAEFHFDQLEDVFLMPKAGSL 416
 DB 285 SRLARVCKKQVGGKNLLAHMMATYLRKALNCISGE--PFYFNEIQSYQL--PSDKS- 339
 QY 417 EYVALSTVSAVFOGFAVCVYHMAIWEVFNPGFAHRDGFQHWMP-YGSKVFPFRGVC 475
 DB 340 REFATPTTSTINGLIGSASFHINETIOAFNGKFEKQSSNSAMPLVLSNRVPERPGTC 399
 QY 476 PSKMTAQPGPFESTKDYDPEVLOFARHPLMWPRPRGRVLYKTHLAQOLHOIIVD 535
 DB 400 -----VNDISLDPDYVNTFIRSHPLMDKAVNHEINNIVYKRDLY--FTKLIVD 446
 QY 536 ---RVEAEDGTVDVIFLGTDGSLKAYIALQAGSAPDEEVELELOVFKPR--TPITENE 691
 DB 447 KIRIDILNQEYIYYVGTMLGRYIKIVQYRRNG-----ESLSKLIDIFEVAPREALQWNE 501
 QY 592 ISVKRQMLVGSRLGVAQRLHOCETGYTACAECCCLARDPYCAMD--GASCTHYRPSLCK 649
 DB 502 ISQTRKSLTIGDHRKQIDILAMCNRRYDNCFRG--VRDPYCGMDKEANTCRPEYELDI-- 557
 QY 650 RFRFRDRIHGNPALQCLQOSEEAVGLVATMYYCTEHNSTFLECLRSPQA-----AV 705
 DB 558 -----LQDVANETSDICDSSVLLKKIYVITGOSVHLGCFVYKIPDEVLMNEQV 603
 QY 706 RWLLQRPQDEGPDQYK--TDERVLTHERGLFRRLSRPDAGTYTC 748
 DB 604 TW-YHHSKDKRGYETIRYPTKYIETTERGLVYVSVNEADGGRYDC 647

RESULT 14
 B49423
 semaphorin I - fruit fly (*Drosophila melanogaster*) (fragment)
 C:Species: *Drosophila melanogaster*
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
 C:Accession: B49423
 R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
 A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A:Reference number: A49423; MUID:94094332
 A:Accession: B49423
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-656 <KOL>
 A:Cross-references: GB:I26082
 C:Genetics:
 A:Gene: semat
 A:Cross-references: FlyBase:FBgn0011259

Query Match 12.7%; Score 601.5; DB 2; Length 656;
 Best Local Similarity 29.2%; Pred. No. 3.9e-38;
 Matches 190; Conservative 94; Mismatches 272; Indels 95; Gaps 24;

QY 208 ECANFVRLQPHNRTHILACGTGAFOPTCALITVGRGHEVHLHPGSVESGRGRCPHEP 267
 DB 3 DCONYIRIMVPSRGLRFLVCGTNSFRPMCTNYIISD-SNTYLE-----ATKQACVCPDP 57
 QY 268 SRPFASPTFIDELVYGLADFLGREAMIFRSGRPPALRSDSGLHDPFRVMAARIPE 327
 DB 58 RHNSISVLADNELYSGLVADSGSDPIIYR--EPLOTQDYDSLISLNAFNV----- 106
 QY 328 NSDDNDKVFEEFSET--VPSPDGGSNHYTVSRVRCVNDVAGGQRYLVNKMSTFLKARLY 386
 DB 107 SSFTQGDVEVYFFRETAVEF INCCK--AIYSRVARVCKMDKGGHRRRNKRTSLKSLN 164
 QY 387 CSVPFGGAEFHFDQLEDVFLMPKAGSLF---VVALFSTVSAVFOGFAVCVYHMAI 443
 DB 165 CSI--PDGYPRYFNEIQSASNLVEGQYSMSKLIYGFNTPSNSIPGSAVCARALDIA 222
 QY 444 EVENGPFARHBDGFQHWMPY-GGKVPFRPGVCPSKMTAQGRFGSTKDYDPEVLOPAR 502
 DB 223 DTFEGQFKEDGTGINSNMLPVNNAKVPDPREGSC-----HNSRALPDPTLNFIR 271
 QY 503 AHPIMFVPRVRHGRPVLYKTHLAQOLHOIIVD-RVEAEDG-TYDVIIFLGTDGSLKVI 560
 DB 272 THSLMDENVAFAFESQPIVLRSTIYRFTQIAYDAQITPBGKTYDVIYFVGTDRKIIKSV 331
 QY 561 ALQAGSAEP-EEVYLELOVFKVPYPTITMETISVKRQMLYGSRLG----- 606
 DB 332 NAEASDADVTSVLEIDVLTKESEIRMLEI--VFTMQYDOPKDSYDDGKLIIVTDS 389
 QY 607 -VAQLRLHOCETYG-TACAECCCLARDPYCAMD--GASCTHYRPSLCKRFRFRDRIHGN 661
 DB 390 QVVALQIHRCHNDKITSQSCVALADPYCAMDKIAGKNSHGAERMLEENYEVQNAATGQ 449
 QY 662 PALQCLQOSEEAVGLVATMYYCTEHNSTFLECLRSPQAARWLLQRPQDEGPD-- 718
 DB 450 HAACPSGKINSKDA---NAGECKGFRNDMLDLSRQSKDQEIIDNIDK--NEGQQTSA 504
 QY 719 -----QVKTDERVLTHERGLFRRLSRPDAGTYTCITLHEGFSQTYVRLALYIVASQLD 773
 DB 505 DIINAQYVETILVMVLAGSIFSLVGFFTG-YFGRRCHG-----KDEDD 548
 QY 774 NFPPEPKPEEPARGGLASTPPEKAWKDILOLIGFANLPRVDE--YCERV 822
 DB 549 NLPYDPTEYEYEORONVNSPSSCRIOQEPKL-----LPQVEEYIYABPV 594

RESULT 15

127165
 hypothetical protein Y54E5B.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27165
 R:Lennard, N.
 A:Title: submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20321
 A:Accession: T27165
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-712 <WIL>
 A:Cross-references: EMBL:AL032663; PIDN:CAA21714.1; GSPDB:GN00019; CESP:Y54E5B.1
 A:Experimental source: clone Y54E5B
 C:Genetics:
 A:Gene: CESP:Y54E5B.1
 A:Map position: 1
 A:Introns: 36/1; 70/2; 85/3; 194/2; 221/1; 253/3; 301/1; 352/3; 522/2; 560/1; 599/3;

Query Match 12.6%; Score 599; DB 2; Length 712;
 Best Local Similarity 29.5%; Pred. No. 6.9e-38;
 Matches 196; Conservative 98; Mismatches 254; Indels 116; Gaps 31;
 QY 161 DRLFGLGIDALYSLRLDQAMPDPREVLPPOGORECVRKGRDPLTCANFVRLQPHN 220
 DB 63 DSLVAGANNAVYNNLSLSTLSVN-HKIDMKPAEHIIECTIMGSKS-TDCQNYIRVLARKS 120

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: October 9, 2002, 13:14:58 ; Search time 24 Seconds
(without alignments)
1411.650 Million cell updates/sec

Title: US-09-813-290-2

Perfect score: 4746
Sequence: 1 MACALAGKVFPMGSPVWHRK.....KMKSRVHAENRPREVAT 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2043.5	43.1	775	1	SM3E_MOUSE
2	2032.5	42.8	785	1	SM3E_CHICK
3	2014.5	42.4	775	1	SM3B_HUMAN
4	1862	39.2	772	1	SM3A_CHICK
5	1846	38.9	772	1	SM3A_RAT
6	1843	38.8	772	1	SM3A_MOUSE
7	1840	38.8	778	1	SM3B_MOUSE
8	1830	38.6	771	1	SM3A_HUMAN
9	1790.5	37.7	761	1	SM3D_CHICK
10	1770	37.3	860	1	SM3D_MOUSE
11	1766.5	37.2	764	1	SM22_BRARE
12	1747	36.8	749	1	SM3B_HUMAN
13	1721	36.3	777	1	SM3D_HUMAN
14	1653.5	34.8	778	1	SM3B_MOUSE
15	1652	34.8	785	1	SM3F_HUMAN
16	1643	34.6	785	1	SM3F_MOUSE
17	1568	33.0	751	1	SM3C_MOUSE
18	1561.5	32.9	751	1	SM3C_CHICK
19	1538	32.4	751	1	SM3C_HUMAN
20	884.5	18.6	837	1	SM4G_MOUSE
21	874.5	18.4	861	1	SM4D_MOUSE
22	867.5	18.3	834	1	SM4C_HUMAN
23	862	18.2	834	1	SM4C_MOUSE
24	861	18.1	838	1	SM4B_HUMAN
25	831.5	17.5	782	1	SM4B_MOUSE
26	825	17.4	766	1	SM27_BRARE
27	776	16.4	770	1	SM4F_HUMAN
28	759	15.8	776	1	SM4F_MOUSE
29	750.5	15.8	776	1	SM4F_RAT
30	706.5	14.9	888	1	SM6B_HUMAN
31	698.5	14.7	888	1	SM6B_MOUSE
32	693	14.6	886	1	SM6B_MOUSE
33	682	14.4	888	1	SM6A_MOUSE

34	678.5	14.3	887	1	SM6B_RAT	070141	rattus norv
35	676.5	14.3	760	1	SM4A_MOUSE	062178	mus musculus
36	664	14.0	931	1	SM6C_MOUSE	09wtm3	mus musculus
37	646	13.6	930	1	SM6C_HUMAN	09ht32	homo sapien
38	645	13.6	960	1	SM6C_RAT	09wt13	rattus norv
39	628	13.2	730	1	SM1A_SCHAM	026473	schistocerc
40	627.5	13.2	771	1	SM1A_DROME	024322	drosophila
41	620.5	13.1	697	1	SM2A_SCHGR	09x268	schistocerc
42	599	12.6	712	1	SM1A_GAEEL	017330	caenorhabdi
43	588	12.4	706	1	SM2A_DROME	024323	drosophila
44	571.5	12.0	1093	1	SM5B_MOUSE	060519	mus musculus
45	555.5	11.7	1077	1	SM5A_MOUSE	062217	mus musculus

ALIGNMENTS

RESULT 1
SM3E_MOUSE STANDARD: PRT: 775 AA.
ID SM3E_MOUSE
AC P70275: 009078: 009079:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3E precursor (Semaphorin H) (Sema H).
GN SEMA3E OR SEMAH OR SEMH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=98175564; PubMed=9515811;
RA Christensen C.R.L., Klingelhoefer J., Tarabykina S., Hultgaard E.F.,
Kramerov D., Lukanidin E.
RT "Transcription of a novel mouse semaphorin gene, M-semah, correlates
with the metastatic ability of mouse tumor cell lines."
RL Cancer Res. 58:1238-1244(1998).
RN [2]
RP REVISIONS.
RA Christensen C.R.L.;
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: MAY BE INVOLVED IN EMBRYONIC DEVELOPMENT.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: DEVELOPING LUNGS, DEVELOPING SKELETAL
ELEMENTS, AND VENTRAL HORNS OF THE DEVELOPING NEURAL TUBE.
CC CORRELATES POSITIVELY WITH TUMOR PROGRESSION.
CC -I- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -I- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: Z80941; CAB02590.1;
CC EMBL: Z93947; CAB07987.1; ALT-SRD.
CC EMBL: Z93948; CAB07988.1; ALT-SRD.
CC MGI:1340034; Sema3e.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003659; PSI.
CC InterPro: IPR001627; Sema.
CC Pfam: PF00047; Ig_1.
CC Pfam: PF01403; Sema_1.
CC SMART: SM00409; Ig_1.
CC SMART: SM00423; PSI_1.
CC Signal: Immunoglobulin domain; Multigene family; Neurogenesis;

KM Developmental protein; glycoprotein.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 775 SEMAPHORIN 3E.
 FT DOMAIN 241 540 SEMA.
 FT DOMAIN 647 736 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 654 729 BY SIMILARITY.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 775 AA: BA41690093D28F40 CRC64;

Query Match 43.1%; Score 2043.5; DB 1; Length 775;
 Best Local Similarity 50.6%; Pred. No. 1,4e-143;
 Matches 398; Conservative 127; Mismatches 235; Indels 27; Gaps 10;

QY MAPSAMALICWLLGGLHGGSSGPPSPGVPRRLSLRYDLSANRSLAFIAGPOGLMIOA 153
 DB 1 MAPAGHILTLTLMLGHLLEMTWTPGHSANPSYPRRLSHKELLELRNRSIFPSLGLFDLHT 60
 QY 154 MYUDEYRDLRFGLGLALYSRLDOAMPDRREVLNPPQPCQRECVKKGDPLETCANFY 213
 DB 61 MLDEYERLFVGGRLVSLNLERVSDGRELTYMSTAVKVECIKMGD-ANECAANYI 119
 QY 214 RYLQPHNRTHLACGTAGOPTCALITYGHRGHEVL-HLEPGSVESGRCRPEPSRPA 272
 DB 120 RYLHNRTHLTLTCATGADPHCAFIRVGHSEPEPLFLEHSHSEGRRCRCPDPNSSEY 179
 QY 273 SFFIOELKTLGLADPLGDEAMFRRSGGPRALRSD-SQSLIHDPFYMAARIENSDQ 331
 DB 180 STLVENELAGLISDYGSDSAIFRSMGKLGIRTEHDEERLKEKFGVSYWIPNEDR 239
 QY 332 DNDKYFEFSETPVSPDGSSNHTVTSRVGVVCVNDAGGQRLVKNKSTFLKARLVSVG 391
 DB 240 DNMKWFETFEKALEENNA-HITYRVRGLCVNDMGGRILVKNKSTFLKARLVSVG 298
 QY 392 PGAEHHPQLEADVLLMKAKGSLVYALFESTVAVFOGFAVCYVHMADIWEVNGRPA 451
 DB 299 MNGIDTYDELDVLLPRDRPKNPVIFGLFNTTINIFRSHVAVCYVHMSISEAFNGPYA 358
 QY 452 HRDQHOHMGPGYVPRPGVPCSKMTAOGPRPGSKDYPRDELOARAHNPMFWV 511
 DB 359 HKEGPEHMSLEKGVYPRPGSCASKV---GKYGTTKDDPDDATIRARIDPLMYQPI 415
 QY 512 RPRHRPVLVYTHLAQQLHQLVYRVEAEDGYDVIPLGTDGSVLYKVALQAGSAEPE 571
 DB 416 KVRHKKPLLVKTDGKYNLRQLAVDRVEAEDGQYDLFICTDGIYLVKITYIVNGETEMNE 475
 QY 572 EVLLELQVFPRTITMEISVKRMLYVSGRLGVALRLHOCETYGACAECLARDP 631
 DB 476 EVLLELQIFNDPAIIMETISSKKQOYLISASAVAOVRFHHCIMYSACADCLARDP 535
 QY 632 YCANDGASCTHRRP-SLGRKRRFRDRIHGNPALQCLGSOEEBAVGLVATWYGTEN 689
 DB 536 YCANDGASCTHRRPFGAHRKRRFRDRIHGNPALQCLGSOEEBAVGLVATWYGTEN 689
 QY 690 NSTFLECPKSPQOAVRLLRPGGEGDQYKTDPRVLTHERGLLRSLRFDAGTYTCT 749;
 DB 596 NSTFLECPKSPQOAVRLLRPGGEGDQYKTDPRVLTHERGLLRSLRFDAGTYTCT 749;
 QY 750 TLEHGSQTVRLVLLVVASQDLNLPPEPRPE-----PPARGLASLPPKAWYDI 803
 DB 656 TVEHNFVYTKITLLEVVEEHKVGMPHFKHDEERHNHMPRLPSGSGQTKP-WKEEF 713
 QY 804 LQILGFANLRVDEYGCRCWCGTTECSGCRSRKQARQKSWAGLELQKKMSRYHA 863
 DB 714 LQILGFANLRVDEYGCRCWCGTTECSGCRSRKQARQKSWAGLELQKKMSRYHA 863
 QY 864 EHNTPR 870
 DB 764 EHNTPR 770

RESULT 2
 SM3E_CHICK
 ID SM3E_CHICK STANDARD; PRT; 785 AA.
 AC 04223; 090666;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3E precursor (Collapsin-5) (COLL-5).
 GN SEMA3E OR COLL5
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_Taxid=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain.
 RX MEDLINE=97470887; PubMed=9331347;
 RA Feiner L., Koppel A.M., Kobayashi H., Raper J.A.:
 RT "Secreted chick semaphorins bind recombinant neuropilin with similar
 RL affinities but bind different subsets of neurons in situ."
 RN Neuron 19:539-545(1997).
 RP SEQUENCE OF 244-543 FROM N.A.
 RX MEDLINE=95329269; PubMed=7605628;
 RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.:
 RT "A family of molecules related to collapsin in the embryonic chick
 RL nervous system."
 RN Neuron 14:1131-1140(1995).
 CC -1- FUNCTION: INDICES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONNS. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
 CC NEURONAL POPULATIONS. BINDS TO NEUROFILIN.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: COLLAPLIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF022947; AAB80952.1;
 CC EMBL: U28243; AAB86899.1;
 DR Interpro: IPR003659; PST.
 DR Interpro: IPR001627; SEMA.
 DR Pfam: PF01403; SEMA; 1.
 DR SMART: SM00423; PSI; 1.
 KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW developmental protein; Glycoprotein.
 FT SIGNAL 1 25
 FT CHAIN 26 785
 FT DOMAIN 244 544 SEMAPHORIN 3E.
 FT DOMAIN 651 740 SEMA.
 FT DOMAIN 741 780 ARG/LYS-RICH (BASIC).
 FT DISULFID 658 733 BY SIMILARITY.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 248 248 N -> D (IN REF. 2).
 FT CONFLICT 250 250 V -> I (IN REF. 2).
 FT CONFLICT 250 250 L -> F (IN REF. 2).
 SQ SEQUENCE 785 AA: 90978 MM: E51BFB717650632 CRC64;

QY 332 DNDKVFEESETPSDGSGNHVTSRVGRVCVNDAGORVLVNMKSTFLKARLVCSVPG 391
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL: U02528; AAC59638.1; -
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR001627; SEMA.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF01403; SEMA; 1.
 DR SMART: SM00409; Ig_1.
 DR SMART: SM00423; PSI; 1.
 KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 22
 FT CHAIN 23 772
 FT DOMAIN 240 538
 FT DOMAIN 643 730
 FT DOMAIN 728 770
 FT DISULFID 650 723
 FT CARBOHYD 53 53
 FT CARBOHYD 125 125
 FT CARBOHYD 591 591
 SQ SEQUENCE 772 AA; 88867 MW; E91E09DE0CC940AC CRC64;
 QY 805 QLTGFANLPRVDEYCEYRMTGTEGSCFRRSRGRKQARGKSMAGELGKMKMSVVAHE 864
 DB 715 QLTGYSNFGQREYCEYRMTGTEGSCFRRSRGRKQARGKSMAGELGKMKMSVVAHE 864
 QY 865 HNRTPR 870
 DB 765 HYRLPR 770

RESULT 4
 ID SMA3A_CHICK STANDARD; PRT; 772 AA.
 AC 090607;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3A precursor (collapsin-1) (COLL-1).
 GN SEMA3A OR COLL1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
 RP 666-680.
 RC TISSUE=Brain;
 RX MEDLINE=94006554; PubMed=8402908;
 RA Luo Y., Raible D., Raper J.A.;
 RT "Collapsin: a protein in brain that induces the collapse and paralysis
 RT of neuronal growth cones";
 RL Cell 75:217-227(1993).
 CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROPTILIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
 CC AND MUSCLE, MODERATE LEVELS IN LUNG, BURSAS, AND HEART AND
 CC VIRTUALLY ABSENT IN LIVER. COLLAPASIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -1- DOMAIN: STRONG BINDING TO NEUROPTILIN IS MEDIATED BY THE CARBOXY

Query Match
 Best Local Similarity 46.8%; Score 1862; DB 1; Length 772;
 Matches 370; Conservative 135; Mismatches 245; Indels 40; Gaps 12;
 QY 103 WLIG-----GLLHGSSPSPPSPVRLRLSTRDLISANRSATFLPGQSLNQLQAYL 156
 DB 3 WLIGIALLSGLVLLAGVNCQAHVKNVRLKLSYKEMLESNNINVFNGLANSSSYHTFLL 62
 QY 157 DEYRDLFLGLGDLATLSLRDQAMPDREVLMPQPGQRECVAKKGSDPLTECNFVRYL 216
 DB 63 DEBRSLIYVAKKHITSFNLVNI-KEYQKIYWPVSHSRBECKAKGADILRECFNFKVL 121
 QY 217 QPHNRTHLACGTGAFOPTCALITYVG-HRGEHYLHLPEGVSGRGRCPHPRPFASTF 275
 DB 122 KTYNQTHLYACGAGFHPMCTYIEVSGHPEDNIFRMEDSHFENGSRGKSPYDKLLTASLL 181
 QY 276 IGCGLTGLTADLGLGHEAMFFRSGGPRPARLRS-DSSLHDPFVMAARKPESDQND 334
 DB 182 VDEGLYSTGAADMGDFALFRTLGHHPIRTBOHOSRWLNDPRFISAHLPESDNPEDD 241
 QY 335 KYVFFSEYTPSPDG-GSNHVTSVRGRVCVNDAGORVLVNMKSTFLKARLVCSVPG 393
 DB 242 KITFFFRFN--ALDGEHTGATIHAKIOLKNDPGHRSILVNMKSTFLKARLVCSVPG 299
 QY 394 GAETHEDQLEDVFLMPKAGKSLVYALFSTVSAVPOGFAVCYVHMADIWENGPFAHR 453
 DB 300 GIDTHDELQDVFLMNSKDKPIYGVFTSSNIFGSAVCMVSMIDVRRVFLGFAHR 359
 QY 454 DGRQHWGYYGGVPPRPGVCSKMTAQGRPGSKIDYPPDEVLOPARAHPLMFVVRP 513
 DB 360 DGRNQHWVYGGVPPRPGVCSKMTAQGRPGSKIDYPPDEVLOPARAHPLMFVVRP 415
 QY 514 RGRPVLVTHLAQOLHQAIVYDVEAEADGTVDYIFLTGDSGVLYKVALDAGSAAEPEV 573
 DB 416 INSRPIMITVDVQQTQIVYDVEAEADGTVDYIFLTGDSGVLYKVALDAGSAAEPEV 475
 QY 574 VLEELQVFPVPPITMEISVKRQMLYVSGRLGVAOLRLHQCEYTGACAECCCLARDPYC 633
 DB 476 VLEELQVFPVPPITMEISVKRQMLYVSGRLGVAOLRLHQCEYTGACAECCCLARDPYC 535

DB 476 LLEEMTVEPRETVISAMKISTKOOOLYIGSATVSOQLPLHRCDPVYKACAECLLARDPYC 535
 QY 634 AMGASCTHREPSLGRRRRRODIRGNPALOC-----LGOSOEBAVGLVATWY 684
 DB 536 AMGSSCSSTFPT-ARRRTRQDIRNGDPLTHCSOLOHNDHNSGQTLLEER-----IT 586
 QY 685 YGHNSTFLECPKSPQAVRWMLORPDEGPDQVCTDERVLTHTERGLFRRLSRDAG 744
 DB 587 YGHNSTFLECPKSPQAVRWMLORPDEGPDQVCTDERVLTHTERGLFRRLSRDAG 646
 QY 745 TYTCTLEHGFSTQVVRVLRVIVASOLDNLFPPEPKPEEPARGGLASTPP--KANYKD 802
 DB 647 IYFCHAVEHGFIOITLLKVTLEVIDTDLLEELLHKEEDADASKTKDATNSMTPSQKIWRV 706
 QY 803 ILLOLGANLPREDCECERWCGTTECCSCFSPRSRGKARGSMAGLELGGKMSRVH 862
 DB 707 FMOLINHPNLTNDEFEQVWKDRKO---RRORPANAQVNTKMKHLDENKKGRRRT 762
 QY 863 AEHNRTPREV 872
 DB 763 HEFERAPRSV 772

RESULT 5
 SM3A_RAT
 ID SM3A_RAT STANDARD: PRT: 772 AA.
 AC 063548:
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=H1STAR; TISSUE=Brain;
 RX MEDLINE=97073089; PubMed=8915837;
 RA Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaegen J.;
 RT "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
 RT relationship to developing nerve tracts during neuroembryogenesis.";
 RL J. Comp. Neurol. 375:378-392(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
 CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
 CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
 CC NEUROPILIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
 CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELECEPHALIC
 CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
 CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
 CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
 CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
 CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
 CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
 CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
 CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
 CC SPINAL MOTONEURONS.
 CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X95286; CAA64607.1;
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003658; PSI.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Iq; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Iq; 1.
 DR SMART: SM00423; PSI; 1.
 KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 772
 FT DOMAIN 240 538
 FT DOMAIN 643 730
 FT DOMAIN 728 770
 FT DISULFID 650 723
 FT CARBOHYD 53 53
 FT CARBOHYD 125 125
 FT CARBOHYD 591 591
 SQ SEQUENCE 772 AA; 88808 MW; 240907812FF9FD2 CRC64;

Query Match 38.9%; Score 1846; DB 1; Length 772;
 Best Local Similarity 46.6%; Pred. No. 5.8e-129;
 Matches 366; Conservative 138; Mismatches 248; Indels 34; Gaps 11;
 QY 101 ICWLLGGLLHGSSGSPSPVRLRLRYRDLISANSAIFLGPOGSLNQLQAMLYDEYR 160
 DB 7 IACLFWGLLRLARANYANGKNNVRLKSTYKEMLESNVITFNGLIANSSTHTFLDLDER 66
 QY 161 DRLEFLGLDALYSRLDQAWDPREVIMPPQPGORECEVKGKRDPLTECANFVRYLOPHN 220
 DB 67 SRLVYGADHIFSEFNLVNI-KDFQKIYVPSYTRDECKVMGKIDILKCANFIKYLKAYN 125
 QY 221 RTHLLAGTGAFOPTCALITTYGHNGE-HVLLHLEGSVSSGGRCHESRFPASTFDGE 279
 DB 126 QTHLYACGTGAFNHPTCTYIEVGHNEDEIFKLDOSHFNFGSKSPYDRLKLTASLLDGE 185
 QY 280 LYTGLTADFLEAMIFRSGGPRPALRSDS-DQSLHPPRVMAARIPENSDDNDKYF 338
 DB 186 LYSGLADPFMRGDRFAIFFTIGHNHRIRTEQHSRLNDPRLISALIPESDNPEDDKYF 245
 QY 339 FFSETVSPDG-GSNHVTYSRVGRVCVNDAGQRYLVNKKSTFLKARLVCSYPGCGAET 397
 DB 246 FFREN--AIDESHGSKATHARIGQICKNDFGHRSVLWKTFTFKARLICSYPGCGNDIT 303
 QY 398 HFDQLEDFLLMPKRAKSLLEYVALFSTYSAVFOGRANOVYIMADITWEVFNCGFAHRDQ 457
 DB 304 HFDELQDFELNMSKDPKMPYVGVFTTSNIFKGSACVMSKSDRVFLGYAHRDGPV 363
 QY 458 HOMVYGGKVPFPPRGVCPKTAOPGRPGSTKDYDPDEVLOFARAHPLMEVPPRRGR 517
 DB 364 YQWVYQGRVFRPRGCTPSKTEG---GFDSTKLPDQVITFAFSHPAMNVPPIINR 419
 QY 518 PVLVTHLAQQLQIVDVREAEQTYVIFLGTDSGVLYKVALIQAGSAPBEVYLEE 577
 DB 420 PIMTIDVNYOFTQIVRVYDAEDQYDMFGLTGIVGLYLVKVSYPKKTMDLEEVLEE 479
 QY 578 IQVKEVPPITTEMETSVKRWMLYVSGRPLGVADLRHOCETGTACAECLLARDPYCAMDG 637
 DB 480 MIVFREPTTISAMELSTQOOLYIGTAGVADPLHRCDIYKACAECLLARDPYCAMDG 539
 QY 638 ASCTHYRPSLGRRRRRODIRGNPALOC-----GOSOEBAVGLVATWYCTE 688
 DB 540 SSCSSTFPT-ARRRTRQDIRNGDPLTHCSOLOHNDHNSGSLER-----IYGYE 590
 QY 689 HNSTFLECPKSPQAVRWMLORPDEGPDQVCTDERVLTHTERGLFRRLSRDAGTYTC 748
 DB 591 NSTFLECPKSPQAVRWMLORPDEGPDQVCTDERVLTHTERGLFRRLSRDAGTYTC 650
 QY 749 TLEHGFSTQVVRVLRVIVASOLDNLFPPEPKPEEPARGGLASTPP--KANYKDIIQL 806

Db 540 SSCSRFFPT-AKRRTTRDRIKNDPLTHCSDDQHNDHNSPLSEER-----IIYGE 590

Qy 689 HNSTFLECLPKSPQAAVRLORPGDEGPDYKTDERVLHTERGLLFRRLSFDAGTYTC 748

Db 591 NSTFLECLPKSPQAAVRLORPGDEGPDYKTDERVLHTERGLLFRRLSFDAGTYTC 650

Qy 749 TLEHFSQTVVRLALVYVASOLDLFPPEPKPEEPARGGLASPP--KMYDIDL 806

Db 651 HAHHEFMOTLLKLVLEIDTEHLEHLKDDGSGSKTKEMSSKTPSQKYVYDFML 710

Qy 807 IGFANPRVDECEWYWGCTECSCFERSRSGKQARGKSWAGLELGRKMSRYVAHEHN 866

Db 711 INPNLNTDEFEQYWKXDRKO---RRQPGHSGSGSSNMKMKHMGESKGRNRRTHERE 766

Qy 867 RTRPREV 872

Db 767 RADRSV 772

RESULT 7

S21B_BRARE STANDARD; PRT: 778 AA.

ID S21B_BRARE

AC Q9W686;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Semaphorin 21B precursor (Semaphorin 1B) (Sema-21B).

GN Sema21B OR Sema3AB.

OS Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OX NCBI_TaxID=7955.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99425174; PubMed=10495275;

RA Roos M., Schachner M., Bernhardt R.R.;

RT "Zebrafish semaphorin 21b inhibits growing motor axons in vivo.";

RL Mech. Dev. 87:103-117(1999)

CC -1- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.

CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY VENTRALLY EXTENDING MOTOR AXONS.

CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC

DR EMBL: AF083382; AAC28103.1; -

DR ZFIN: ZDB-GENE-991209-6; Sema2ab.

DR InterPro: IPR003599; 19.

DR InterPro: IPR003006; 19_MHC.

DR InterPro: IPR003659; PSI.

DR InterPro: IPR001627; Sema.

DR Pfam: PF00047; Ig_1.

DR Pfam: PF01403; Sema; 1.

DR SMART: SM00409; IG; 1.

DR SMART: SM00423; PSI; 1.

KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.

FT SIGNAL 1 17

FT CHAIN 18 778

FT DOMAIN 241 539

SEMA.

FT DOMAIN 645 723

FT DOMAIN 721 776

FT DISULFID 652 716

FT CARBOHYD 54 54

FT CARBOHYD 127 127

FT CARBOHYD 593 593

SO SEQUENCE 778 AA; 88904 MW; 4D36F4332A21895 CRC04;

Query Match 38.8%; Score 1840; DB 1; Length 778;

Best Local Similarity 47.9%; Pred. No. 1,6e-128;

Matches 378; Conservative 123; Mismatches 258; Indels 30; Gaps 11;

Qy 99 WAICWLLGGLL-HGSSGSPGSPGSPRLSTYDLSANSAFLPOGSLNQAMTYD 157

Db 5 WIVILLIWTLLAPERGTVAAORSKSNVPLRKPSYKEMESNMLTFNGIANSAYHTFLD 64

Qy 158 EYRDLFLGIGDALYSLLDQAMPDPPEVLPPOPQREECVRCRPLTCANFVVLQ 217

Db 65 EERGLFVGAKDHLSENLVDINMDQDLISNPSFSRDECKMAGKDYOKCAMPFIVLQ 124

Qy 218 PNRTHLLACGTGAFOPCALITVGRGEHVLHLEPGVESGRCRCHPSRPASTFID 277

Db 125 PENOHLIACGTGAFHPCAHAYGVGRKSEDMTFRLGSEFNGRGRKSPYKLOTASMLID 184

Qy 278 GELVYGLTADFLGRAMIFRSGPRPALRSDS--DSSLHDPRTVMAARIIPNSQDNKV 336

Db 185 GELVYAGTSADPFWGRDFAFLFKLHHPRTEDQHSRWLNDPPEVSLHLPSSDNEEDKI 244

Qy 337 YFFESFTVPSPDGGS-NHVTVSRGVGVNDAGGQGVLVNKSFTLAKRLVCSVPGPGA 395

Db 245 YLFFERN--AIDGEOISKATHRIQOLCKNDGGRSLVNMKTTFLKRLVCSVGLNGI 302

Qy 396 EHFEDOLEDFVLLMPKAGKSLVYALFSTVSAVFGFAVCYTHMADLIEVNGPRAHDG 455

Db 303 DTHFDELQVFLMSSKDPKNPITVAVFTSSNIFGSAVCMTSMADIRVFLGPAHBDG 362

Qy 456 POHOKPRTYGGKVPFRPGVCPSPKMTAGPGRPGSTKVDYDEVLOFAHAPLFWFVPRH 515

Db 363 PNYQWVFLNRPVPRPGVCPSPKTF--DGEFTKDPDDVITTAHSHPMNTVFPIN 418

Qy 516 GRVLVKTLLAQLQIIVDVEAEEDGYVIFLGTSGSVLKVALAOLAGSAPREEVVL 575

Db 419 NHPIITIKTDVDFPQIIVDVEAEEDGYVIFLGTDMGTVLKVASIPRGWHDLEEVLL 478

Qy 576 EELQVKEVPTITEMEISYKRMILYGSRLGYAQLRHOCETTYGTAACACCLARPYCAM 635

Db 479 EEMTVREPAITAMELSTKOOLYLSAIGVQMPDLCRDVYGRACACCLARPYCAM 538

Qy 636 DGASCHYRSLGKRFRFRDRIHGNPDLQCCGQSOEEA---VGLVATVYGTENHST 692

Db 539 DGSQCSKTFPT-AKRRTTRDRIKNDPLTHCSDDQHNDHNSPLSEER--YGVENSSS 596

Qy 693 FLECLPKSPQAAVRLORPGDEGPDYKTDERVLHTERGLLFRRLSFDAGTYTC 752

Db 597 FLECLPKSPQAAVRLORPGDEGPDYKTDERVLHTERGLLFRRLSFDAGTYTC 656

Qy 753 HGSQIVVRLALVYVASOLDLFPPEPKPEEPARGGLASPPKMYDIDL 812

Db 657 HGFLOTLLRLTLVIPAHLDDLHPRDPTDNPAMG-----KMYRDFSLINPPSP 709

Qy 813 PRVDECEWYWGCTECSCFERSRSGKQARGKSWAGLELGRKMSRYVAHEHN 863

Db 710 NSVDQCEQYWKXDRKO---RRQPGHSGSGSSNMKMKHMGESKGRNRRTHERE 769

Qy 864 EHNTPREV 872

Db 770 EMGRAPRSV 778

RESULT 8

SM3A_HUMAN STANDARD; PRT: 771 AA.

ID SM3A_HUMAN

AC Q14563;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=94094332; PubMed=8269517;
 RA Kolodkin A.L., Matches D.J., Goodman C.S.;
 RT "The semaphorin genes encode a family of transmembrane and secreted
 growth cone guidance molecules.";
 RL Cell 75:1389-1399(1993).
 RN [2]
 RP SEQUENCE OF 1-37 FROM N.A.
 RA Woessner J., Malm P., Hinds K., Strommatt C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 39-182 FROM N.A.
 RA Rohlfing T., Tin-Mollam A.M., Duckels G.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INDICES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
 CC NEUROFILIN-1/PLEKXIN-1 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMOBILIZIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: L26081; AAA65938.1;
 DR EMBL: AC004451; AAC06185.1;
 DR EMBL: AC004848; AAC78622.1;
 DR MIM: 603961;
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Iq; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Iq; 1.
 DR SMART: SM00423; PSI; 1.
 DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 771 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 642 729 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 727 769 ARG/LYS-RICH (BASIC).
 FT DISULFID 649 722 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 771 AA: 88889 MW: 9985f8D3EAED8456 CXC64;
 Query Match 38.6%; Score 1830; DB 1; Length 771;
 Best Local Similarity 46.5%; Pred. No. 8,8e-128;
 Matches 366; Conservative 139; Mismatches 244; Indels 38; Gaps 13;

QY 101 ICWLLGGLLHGSSGSPSPSVPRRLRLSYRDLLSANRSATFLGPGQSLNLAQWLYDEXR 160
 DB 8 VC-LFWGVLLTARANYQNGKNVPRKLSTYKEMLSNNVITFNGLANSSSYTFLDLDER 66
 QY 161 DRLFFGGIDALYSRLDQAWPDRREVLPPOPQGECEVCRKGRDPLTCANFVRYLQPHN 220
 DB 67 SRLVGAADHIFSEPLVNI-KDQKIWPVSVYTRDECKKMGKDLKCANFIKVLKAYN 125
 QY 221 RTHLLACGTGAFOPTCALITVGNRGE-HVLLLEPGSVSGRCRCPHSPFASFFIDGE 279
 DB 126 QTHLYACGTGAFHPICTYIEIGHPEDEIFLKNHSEFNGSKSPYDKLLTASLLIDGE 185
 QY 280 LYTGTATFLREAMIFRSGGPRRLASDS-DQSLHDPRFVMAARIPENSDDNDKYF 338
 DB 186 LYSGADEWGRDAIFETLGHHPDIRTEQDHSRLNDPKFISALLISESDNEDDKYF 245
 QY 339 FFESEVSPDG-GSNHVSVSRGVCVADGQAVLVKMSFTFLKARLYCVSPRGAGET 397
 DB 246 FFEREN-AIDEGHGKATHARIGQICKNDFGHSVLNKKMTFLKARLYCVSPRGAGET 303
 QY 398 HFDOLEDFLLMPKAGSLLEVYALFSTVSAVGGFVAVCYHMAIDWIEFNGPFARHDPQ 457
 DB 304 HFDELQVFLNFKDPKPNPVYGVFTTSSNIFKGSACVMSMSDVRFLGPAHRRGN 363
 QY 458 HMGPRYGVKPEPRPGVCPKMTQPGRPSTQDYDEVLYQFARHPLMFWPRPHGR 517
 DB 364 YQWVPYQGRVPYPRPGTCSPTGS-----GDSYKDLDPDVTTFARSHPMVNPVPMNMR 419
 QY 518 PVLVKTLLAQQLQIVVDVVEADGTYVIFLGTDSGVLYKVLQAGSGAPEEVLVEE 577
 DB 420 PIVKIVDQVNYQFQIVVDVVEADGQYVMEIGTVGTVLKYVISPREWYDLEVLLEE 479
 QY 578 LOYEKVPPTTEMEISKRMQLVYVSGRLGVAQLRHOCETTYGTACACCLARDPYCAMDG 637
 DB 480 MYFREFETALISAMELSTKQQLYIGTAGVAQLPLHRCIDYGRACACCLARDPYCAMDG 539
 QY 638 ASCTHRRPSLGRKRRFRQDRIHGNPALQCL-----GQSOEEAVALVATMVGTGH 669
 DB 540 SACSRYEPT-AKRRTRQDRIHGNPALQCL-----GQSOEEAVALVATMVGTGH 669
 QY 690 NSPFLFELKPSQAAVWMLQRPDGPDOYKTDERYLTERGLFRRLSRFAGTYTCT 749
 DB 591 SSTFLFELKPSQAAVWMLQRPDGPDOYKTDERYLTERGLFRRLSRFAGTYTCT 749
 QY 750 TLEHGFSGQVRLALVIVASQDNLFPPEKPEDEPARGLASTPP--KAWKDLQL 807
 DB 651 AVDHGFIQTLKXVLEVIDENLELLHKDDGDSKTKMSMSMTPSQKWTYRDMQL 710
 QY 808 GFANLPRVDEYCEWVRCGTTECGSGCFRSGRGAQGRKS--WAGLELKKMKSRVAEH 865
 DB 711 NHENLMTDEFCBQVWKRDRKQ-----RRPRGHTPGNSNKKHLEQENKKGNRRTHFE 764
 QY 866 NRTPREV 872
 DB 765 ERAPRSV 771
 RESULT 9
 SM3D_CHICK STANDARD; PRT: 761 AA.
 AC 090663;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3D precursor (Collapsin-2) (COLL-2).
 GN SEMA3D OR COLL2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; UZ8240; AAA86896.1; -
DR	InterPro: IPR003006; IG_MHC.
DR	InterPro: IPR003600; IG_Like.
DR	InterPro: IPR003659; PSI.
DR	InterPro: IPR001627; Sema.
DR	Pfam: PF00047; Ig_1.
DR	Pfam: PF01403; Sema; 1.
DR	SMART; SM00410; IG_Like; 1.
DR	SMART; SM00423; PSI; 1.
KW	Signal; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.
FT	SIGNAL 1 24 POTENTIAL.
FT	CHAIN 25 761 SEMAPHORIN 3D.
FT	DOMAIN 245 543 SEM.
FT	DOMAIN 646 726 IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN 727 757 ARG/LYS-RICH (BASIC).
FT	DISEULFID 653 719 BY SIMILARITY.
FT	CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 761 AA; 87300 MW; 3E09AE3DBA55F46B CRC64;

```

QY      93  SMASNAICWLLGGLHGGSSPSGSPRYRLYSRLLSANSRAIPAGGGSJNLO 152
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4  SÖVRNA---CGLSLAMIEFFRYGTST-QNIPRIKLSYKDLLSNSIPLPGSEIGDER 59

QY      153 AMYIDEXYRLEFLGGLDALYSLRIDQAMPREVLWMPORQOREECYKRCRDPITECANF 212
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      60  TLLDEERGRLLYGAKDHIFELLYLVNDLNKMYKCIYMPAAKEMKELCIACKRADHTDCAE 119

QY      213 VRVLDQPHNRHTLLACGTCGAPRPTCALITYG-HREHYNLHLEPGSEVSGRCGRPEPRPF 271
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      120  IRLVLDQPHNRHTVLYVYCGTGAFHPLCGYTELGHNKKEITFLRDTQWLESGRLCRPPDOOF 179

QY      272  ASTFIDELGYTLGVLADFLGREAMIFRSGGP--RPALRSR--SDQSLHLDRPFVMAARIPE 327
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      180  ASYMADELYLXGCTASDFGLRGTALTIRSGRSHDNHVIIRTOISEHYWYLTGAKFIATPIPD 239

QY      328  NSDQDNKVYFFESFETVSPDGGSNHVTY-SRQRYCVYNDAGGRVLYNKKSTYFKARLY 388
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      240  TYNPDQDITYEFFREI--SODSSTSDKTIILSRVGRVCCNDMGGRSLINKWTTLFKALY 299

QY      387  CAVPRGGAETHFQGLDEVLFLMPRAGSLSEVYALLFSTVSVAPOGFACVYHMADIWEVF 444
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

	RESULT 10	
ID	SZ1A-BRARE	PRT; 860 AA.
AC	Q9W7J1;	STANDARD;
DT	30-MAY-2000 (Rel. 39, Created)	
DT	30-MAY-2000 (Rel. 39, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Semaphorin Z1A precursor (Semaphorin 1A) (Sema-Z1A).	
GN	SEMAZ1A OR SEMAZ3A.	
OS	Brachydanio rerio (zebrafish) (zebra danio).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;	
OC	Cypriniformes; Cyprinidae; Danio.	
OX	NCBI_TaxID=7955;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
KC	TISSUE=Embryo;	
RX	MEDLINE=99313409; PubMed=10386838;	
RA	Yee C.S., Chandrasekhar A., Halloran M.C., Shoji W., Warren J.T.,	
RA	Kuwada J.Y.;	
RT	"Molecular cloning, expression, and activity of zebrafish semaphorin	
RT	Z1a.";	
RL	Brain Res. Bull. 48:581-593(1999).	
CC	-I- FUNCTION: MAY INFLUENCE OUTGROWTH BY A VARIETY OF GROWTH CONES	
CC	INCLUDING THOSE OF THE POSTERIOR LATERAL LINE GANGLION.	
CC	-I- SUBCELLULAR LOCATION: Secreted (by similarity).	
CC	-I- DEVELOPMENTAL STAGE: EXPRESSED IN HIGHLY SPECIFIC PATTERNS WITHIN	
CC	THE DEVELOPING EMBRYO.	
CC	-I- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.	
CC	-I- SIMILARITY: CONTAINS 1 SEMA DOMAIN.	
CC	-I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/)	


```

CC or send an email to license@lsb-sib.ch).
CC EMBL: AF086761; AADA3964.1;
DR ZFIN: ZDB-GENE-991209-3; sema3aa.
DR InterPro: IPR003599; 1g_1MC.
DR InterPro: IPR003006; 1g_1MC.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; 1g_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00423; PSI; 1.
KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 860
FT DOMAIN 241 539
FT DOMAIN 645 724
FT DOMAIN 722 858
FT DISULFID 652 717
FT CARBOHYD 53 53
FT CARBOHYD 126 126
FT CARBOHYD 593 593
SQ SEQUENCE 860 AA; 97263 MW; 5FD4C12194F5165C CRC64;

Query Match 37.3%; Score 1770; DB 1; Length 860;
Best Local Similarity 44.9%; Pred. No. 2.8e-123;
Matches 347; Conservative 144; Mismatches 245; Indels 36; Gaps 10;

OY 104 LGLGLHGSSGSPSPSVRLRLSLYRDLISANRSLAFLCPQSLNLAQMYLDEYRDL 163
DB 10 LCCVALPGRVAPDHTENVRKLTSLNEMLESSNLVFTGLANSSGYDFRLMGERL 69
OY 164 FLGLDALYSLRLDQAMPDREVLPPQPGRECVKRGKRDPTFCANFVVLDPHNHT 223
DB 70 LVGAEHVEFSDLYNINRDYQIAWPAFPRKDECKWAKDLRDCSNFVAVLSYDTH 129
OY 224 LMGCGAPFTCALIVGRGE-HVHLEPGSVESGRCPHEPSPREASTFIDELTYT 282
DB 130 IYCGTGFAPFISFELMGRKRAEDNIFRLDANYFENGRKSPYDPKKQSSLLDGEYS 189
OY 283 GLTADFLGRAMIFRSQGPRLALRSDS-DQSLDHPREVAARIPENSDDNDKVFEEFS 341
DB 190 GTSADFGGRDFAIFRLTGHHPRTEDQHSRWLNEPFLGHLPESDNEDDKIFLFFK 249
OY 342 EYVPSDPG-GSNHYTVSRVRCVNDAGGRVLVNKKSTFLKALVCSVPGGAEETHFD 400
DB 250 EN--AMDGHTGKATISRIQOLCKNDMGCHRSIVNKKWTFLKAKLTGSPGLNIDTHFD 307
OY 401 QLEDFVFLMPKACKSLLEVVALFSTVSAVFGGFVAVCYHMDINEVFNPPRAHNDGPOHOW 460
DB 308 ELQDVLMSKDKRNKYITAVFTSSNIFRGSALICMSMADIRRVFLGPRYAHNRGPRYQW 367
OY 461 GRYGKAVPPRPVCPSCSKMTAOPGRFPGSTKDYDPDEVLQPARAHPLMFVWPRRGRPV 520
DB 368 VPFQGHVPRPRPCTCPSKTFG---GFDSTKDLDPDYITFARLHPANYNVQMGKRPV 423
OY 521 VKTHLAQQLQIVVDVEADGTYVIFLGTSQSVLKVITAIQAAGSADEEVLEIYQ 580
DB 424 VRTNVEYQITQLVVDVEADGQYDVFETGDTGLVTKVITPRESWHLDEEVLEEMTV 483
OY 581 FVVPFTFMEISVKRQMLYVGSRLVGAOLRLHOCGYGACAECLARDPYCAMPGASC 640
DB 484 FPEPFTITMELSTKOQQLYLSDLGISOPLHRCVEYGAACECLARPYCAMPGATEC 543
OY 641 THYRSLGRARRRRDIRGNPALOC-LGQSQEEAVGLVATATWYTGTEHNSTFELCP 698
DB 544 SYTFPT-ARRRRRDIRGDDPLSGCSDLHNDLDEGYSVEERSVYGVENSMTFELCP 602
OY 699 KSPQAAVRLRLQPGEGEDQYKTBERVLTHERGLFRLSRFDAGTYCTTLEHFSQT 758
DB 603 KSQALLITWQLKPYDERKHEIVIDERLSLTQGLIRSLTQADSGVFLCHAVEHGFIQ 662

```

```

OY 759 VVRLALVIVASQDLNLEPPEKPEEPARGGLASTPPKAWYDILQIGFANLPRVDEY 818
DB 663 LRRINQVIPSQVSGELLRLRAGINDKDP-----PKHLMWRDFPSLLEHNDLVNDEF 716
OY 819 CRRVNCRGTTESGCRSRSGKQARGKSMAGLICKKRSRYVAEHNTPR 870
DB 717 CERIM-----KREKKPRGKRAKRVNPG---TGVSIKNEKTPQ 750

RESULT 11
SM22_BRARE
ID SM22_BRARE STANDARD; PRT; 764 AA.
AC Q9WGG6;
DI 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Semaphorin 22 precursor (Semaphorin 2) (Sema-22).
GN SEMA22 OR SEMA2.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99112778; PubMed=9915572;
RA Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Raper J.A.,
RA Kuwada J.Y.;
RT "Analysis of a zebrafish semaphorin reveals potential functions in
RT vivo."
RL Dev. Dyn. 214:13-25(1999).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE GUIDANCE OF SEVERAL AXON
CC PATHWAYS.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN A DYNAMIC AND RESTRICTED PATTERN
CC DURING THE PERIOD OF AXON OUTGROWTH.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
DR EMBL: AF124485; AAD21310.1; -.
DR ZFIN: ZDB-GENE-990715-2; sema2.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 41
FT CHAIN 42 764
FT DOMAIN 261 559
FT DOMAIN 661 740
FT DISULFID 741 762
FT DISULFID 668 733
FT CARBOHYD 143 143
FT CARBOHYD 490 490
FT CARBOHYD 610 610
SQ SEQUENCE 764 AA; 87859 MW; A3ED95C2C479D7AE CRC64;

Query Match 37.2%; Score 1766.5; DB 1; Length 764;
Best Local Similarity 45.3%; Pred. No. 4.4e-123;
Matches 359; Conservative 137; Mismatches 245; Indels 52; Gaps 18;

OY 84 RRQRCQPFSSMAPSAWALWLAGGLLHGSSGSPSPSVRLRLSLYRDLISANRSLAFL 143
DB 111

```

```

Db 10 RRORQVYRGFRSCAM---WSTSVNLFPSLPEGNCKMESLPRVKGKYLHRSVYFET 66
QY 144 GPGGSLNLOAMYDEYDRFLFLGGLALYSRLDOAMPDREVIAMPPOGQRECVKGR 203
Db 67 GSSEGHQFQVLLDEEERSRLLGAKDHVYLLDPNINKHPRKLSWPARSDREVCEILAGK 126
QY 204 DPLTECANFVYVLOPHNRTHLLACGTAFOPTCALITY-GHGRGHVLLFEGSVESGR 262
Db 127 NPLTECANFVYVLOPHNRTHLLACGTAFOPTCALITY-GHGRGHVLLFEGSVESGR 186
QY 263 CPHEPRPFASTFIDELTYGLTADFLGREAMIFRSGGPRP---ALNSD-SDQSLHDP 318
Db 187 CFPDPNPFASVLTLDYVLAGTASDFLGTSTFRSLGCPPHQOYIRITDISEDYINEGK 246
QY 319 FVMAARIPENSDDNDKVYFFEFSTVSPDGGSHVY-SVGRVCVNDAGGVVYVWKK 377
Db 247 FISHPISDYNPDOKITFEFFRBA--SRDSTYDKSVSHVARIKRDVAGLSLNTKW 304
QY 378 STFKARLVCSVPGGAETHEFDLEDFLLMPKAKSLLEYALFSTVSAVFOGFAVCY 437
Db 305 TTFKALVCSIPGPDVDTHFDELQDFLLPSRDEKNPMYGVFTTSSIFKSAVCY 364
QY 438 HMADIWEFNGPFAHRCGPOHMGVPGKVPFPRGVCSPKTAQPGREGSTYDYPDEV 497
Db 365 TMEDIAFAFNGPYAHKRGPDHRRWEYEGRIPIYPRGTCPSR-TYDP--HIKTKDFPDEV 421
QY 498 LQFARHPFMPVPRPRIGRVLYKTHLAQOLHIVDRAVEDGTIVITLGDSSVL 557
Db 422 ISFRLHPLAMQSVHAPMTGRIFTRINTEYRLTOIVRAVEDQVAMFLGTDMSVL 481
QY 558 KVAL-QAGSGAEPEEVYLELOVFKVPTITEMETSVKROMLYVSLGVAQLRHOCE 616
Db 482 KVSITQENMS--EELTELOVFKNSPLINMEVSKOOLPFGGSDGLVQVSLHNCQ 539
QY 617 TYGACACCLARDPYCAMDCACTHYRPSLGRKRRRQDRIHGNPAIQ-----LQOS 670
Db 540 IYGGGACACCLARDPYCAMDCTGSRYP--SKRRARODIKHGPSSHCWMTEDVLRN 598
QY 671 QEEENVGLVATVYVGHNSFLECLKPSQAARWMLDORPGDGPPOVKDEKVLHTE 730
Db 599 VEK-----VLTVGSNSFLECVSKSOALIRWYLVKRGVHROEIKRDEKVLITD 650
QY 731 RGLFRRLSRPDATYCTTLEHGFSGQTVRLAVLVIASOLDLFPPEKPEPPARAG 790
Db 651 RGLFRRLSRPDATYCTTLEHGFSGQTVRLAVLVIASOLDLFPPEKPEPPARAG 706
QY 791 LASTPRKAMYKDILDLIGFANLPYDEYCEWCGTETGSCCFSRSGKARCKSMAG 850
Db 707 AVTEPRORYKDYLRML-SGPARSIDECETMW-----HREKKOKRGR-WKH 751
QY 851 L-ELGKKMKSRVH 862
Db 752 VOELKRSRRRRH 764

```

```

RT "Human semaphorins A(V) and IV reside in the 3p21.3 small lung
RT cancer deletion region and demonstrate distinct expression patterns.",
RT Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Dante M., Wamsley P.
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFIC TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
CC SIMILARITY)
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
CC ENDOPLASTIC RETICULUM.
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
CC VARIETY OF NEURAL AND NONNEURAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28369; AAD09138.1; -
DR EMBL: U73167; AAC02731.1; -
DR MIM: 601281; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00423; PSI; 1.
KW Signal, Immunoglobulin domain: Multigene family; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 749
FT DOMAIN 239 537
FT DOMAIN 637 717
FT DOMAIN 699 702
FT DOMAIN 724 744
FT DISULFD 644 710
FT CARBOHYD 82 82
FT CARBOHYD 124 124
FT CARBOHYD 427 427
FT CONFLICT 29 36
FT SEQUENCE 749 AA; 83121 MW; 1F3B8F63F59444F3 CRC64;
SQ
Query Match 36.8%; Score 1747; DB 1; Length 749;
Best Local Similarity 47.1%; Pred. No. 1,2e-121;
Matches 367; Conservative 107; Mismatches 239; Indels 66; Gaps 17;

```

Db 219 WLNPKVVKVFWIESENPDDDKIY FFEFRETAVEAARALGLSVSRVGOICRNDVGGORS 278
 QY 373 LVNKMSTFLKARLVCSYVPGGAETHFDOLBDVFLMLPKACKSLVLAFFSTVASVPOGF 432
 Db 279 LVNKMSTFLKARLVCSYVPGGAETHFDOLBDVFLMLPKACKSLVLAFFSTVASVPOGF 337
 QY 433 AVCYVHMADWVEVNGPRARHDSQHOHMGYGVKVPFPGVCSKMTADGPRFGSTKD 492
 Db 338 AVCYVHMADWVEVNGPRARHDSQHOHMGYGVKVPFPGVCSKMTADGPRFGSTKD 393
 QY 493 YPDEVLOFARHAFEMFVNRHGRPRVLYKTHLAQOLHYVDVVEADGTYDYVFLGTD 552
 Db 394 FPDVLOFARHAFEMFVNRHGRPRVLYKTHLAQOLHYVDVVEADGTYDYVFLGTD 453
 QY 553 SGVLYKLVIALQAGSAPBEVEVLEELQVFKVPTTEMEISVKROMLYGVSRGLVAQRL 612
 Db 454 VGVLYKLVIALQAGSAPBEVEVLEELQVFKVPTTEMEISVKROMLYGVSRGLVAQRL 513
 QY 613 HOCETGTACAECCCLADPYCAMDGASCTHYRPSLGRFRFRDRIHGNPALQCLGOSQE 672
 Db 514 HRCAMHGRVCTECCCLADPYCAMDGASCTHYRPSLGRFRFRDRIHGNPALQCLGOSQE 572
 QY 673 EEAVALTAATVMTVGTENHSPFLECLTPKSPAAVRYMLORPEDGPDVKTDERVLTHERG 732
 Db 573 P-----ALLEHKVFEVSGSAFLCEPESLSQARVMTFORAGVTAHTOYLAERTERTARG 628
 QY 733 LLEFRLSRPDAGTYCTTLEHFSQTVRLALVYVVAQDLNLF-PEPKPEEPARGGL 791
 Db 629 LLEFRLSRPDAGTYCTTLEHFSQTVRLALVYVVAQDLNLF-PEPKPEEPARGGL 684
 QY 792 ASPEPKMYWDIOLI-----GFANLPRVDEYCEVRCVGTGTESSGC-FRERSGQAR 844
 Db 685 -----PKMYDFOIOLVPEPGGGGANSILRM-----CRPOLQSLPLESRGRGRNR 731
 RESULT 13
 SM3D_HUMAN STANDARD; PRT; 777 AA.
 AC 095025;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3D precursor.
 GN SEMA3D.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Mead K., Graves T., Wilson C.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INDICES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONNS. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
 CC NEURONAL POPULATIONS. BINDS TO NEUREPHILIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DOMAIN: STRONG BINDING TO NEUREPHILIN IS MEDIATED BY THE CARBOXY-
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AC004957; AAC83081.1; ALT_SEQ.
 DR InterPro: IPR003589; Ig.
 DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003659; PSI.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00403; Ig; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Ig; 1.
 DR SMART: SM00423; PSI; 1.
 DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 39
 FT CHAIN 40 777
 FT DOMAIN 257 555
 FT DOMAIN 658 738
 FT DOMAIN 739 773
 FT DISULFID 665 731
 FT CARBOHYD 139 139
 FT CARBOHYD 607 607
 FT CARBOHYD 724 724
 SQ SEQUENCE 777 AA; 89651 MW; 3F7B0D7AF50F53BD CRC64;

Query Match 36.3%; Score 1721; DB 1; Length 777;
 Best Local Similarity 45.1%; Pred. No. 1e-119;
 Matches 346; Conservative 144; Mismatches 229; Indels 48; Gaps 17;

QY 122 SVPLRLSYRDLSSANRSALFLGQGSINLAQMYLDEYRDLFLGLDALYSLRLDQAMP 181
 Db 41 NIPRLKLYKDLILNSNCIPFGSSEGIDFQTLLEDEGRGLLGAKDHIFLSLVLDNK 100
 QY 182 DPREVLPPOGORECECRKRDPLTECANFVRYLOPHNRHLLACSGAOPICALITY 241
 Db 101 NFKTIYPAKEREYELCKLACKDANTECANFIRVLOPHNKTHIYCGGAGHPICGYIDL 160
 QY 242 G-HRGEVHLLEPGSVESGRGCRPEPSPASTFIDELTYGLTADFLGREAMIFRSGG 300
 Db 161 GYKEDILIFKLDTHNLHSGRLKCPEDPOQPAVSMTDEVILSGASDPLGDTAFNTSLG 220
 QY 301 P---RPLRSD-SQSLIHPRFVMAARIPENSQDNDKYFFSEVFPSPDGSSNHYV 356
 Db 221 PTHDHYIRTDISHYWMNGAKKFTGTFEIPDTYNPDDDKIYFFRES--SQEGSTSDKI 278
 QY 357 -SRVGRVCMVDAGGQRYLVNKMSTFLKARLVCSYVPGGAETHFDOLBDVFLMLPKAGKS 415
 Db 279 LSRGRCKNDVGGQRLINKMTFLKARLVCSYVPGGAETHFDOLBDVFLMLPKAGKS 338
 QY 416 LEVYALFTSVSAVFOGAVCYVHMADWVEVNGPRARHDSQHOHMGYGVKVPFPGVCS 475
 Db 339 PVYGVFTTSSIFKGSVAVCYVHMADWVEVNGPRARHDSQHOHMGYGVKVPFPGVCS 398
 QY 476 PSKMTAQGRPFSGTKDYPDEVLOFARHAFEMFVNRHGRPRVLYKTHLAQOLHYVD 535
 Db 399 PSK-TYDP--LISTRDPDPDVLSFIRHSVMKSVYPVAGGPFERKINVDYRLTQIVD 455
 QY 536 RVEAEDGTVDVIFLGTDSGVLYKLVIALQAGSAP---EEVLEELQVFKVPTTEMEI 592
 Db 456 HVAEDQGYVMFLGTDIGVLYKVS-----SKKMMEEVLEELQVFKVPTTEMEI 511
 QY 593 SVKROMLYVSRGLVADRLHOCETGTACAECCCLADPYCAMDGASCTHYRPSLGRFR 652
 Db 512 SLKQOQLYTVSRGLVADRLHOCETGTACAECCCLADPYCAMDGASCTHYRPSLGRFR 570
 QY 653 RRODIRGNALOC-----LGOSQEEAVGLAATVWYVEHNSSTFLTECPSPQAAVR 706
 Db 571 RRODVAKGDPITQOCWDLSDSHETADEKV-----IFGIERSTFLTECPSPQAAVR 623
 QY 707 WLORPDEGPDVKTDERVLAHTERGLLFRRLSRFDAGTYCTTLEHFSQTVRLALV 766
 Db 624 WYIORSGEDEHEELKPERILITEYGLLIRSLQKDSGMYCAQAEHTFIHTIKVLNV 683
 QY 767 IYASQDLNLFPEPKPEEPARGGLASTPRKAMYKTDLIGRANLPRVDEYCEVRCV 826
 Db 684 IENQOMEN--TORAHEEGVKDLAES--RLRKVDIQLTSSPNF-SLDQYCEOMW--- 735
 QY 827 TTECSGCFRSRSGKQARKSMAGL-ELGKKMSRYVAENRPPREV 872

RP SEQUENCE FROM N.A., AND VARIANT MET-503.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96230324; PubMed=8786119;
 RA Xiang R.-H., Hensel C.H., Garcia D.K., Carlson H.C., Kok K.,
 RA Daly M.C., Kerbacher K., van den Berg A., Veldhuis P., Buys C.H.C.M.,
 RA Naylor S.L.,
 RT "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
 RT 3p21, a region deleted in lung cancer.";
 RL Genomics 32:39-48(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Nelson J., Biewald T.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 394-436 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96210603; PubMed=8633026;
 RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
 RA Albanesi J.P., Lee C.-C., Letman M.I., Minna J.D.,
 RA "Human semaphorin IV) and IV reside in the 3p21.3 small cell lung
 RA cancer deletion region and demonstrate distinct expression patterns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
 CC -1- FUNCTION: MAY PLAY A ROLE IN CELL MOTILITY AND CELL ADHESION.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity)
 CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
 CC VARIETY OF NEURAL AND NONNEURAL TISSUES. THERE IS HIGH EXPRESSION
 CC IN MAMMARY GLAND, KIDNEY, FETAL BRAIN, AND LUNG AND LOWER
 CC EXPRESSION IN HEART AND LIVER.
 CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 10.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in, no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: U33920; AAC50568.1; -;
 DR EMBL: U38276; AAB18276.1; -;
 DR EMBL: AC000063; AAB46344.1; -;
 DR EMBL: U32171; AAB06011.1; -;
 DR EMBL: U32172; AAB06012.1; -;
 DR MIM: 601124; -;
 DR InterPro: IPR003599; Ig_
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SMO0409; Ig; 1.
 DR SMART: SMO0423; PSI; 1.
 KW Signal. Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 785
 FT DOMAIN 272 569
 FT DOMAIN 671 753
 FT DOMAIN 758 779
 FT DISULFID 678 746
 FT CARBOHYD 53 53
 FT CARBOHYD 126 126
 FT VARIANT 474 474
 FT
 FT VARIANT 503 503
 FT
 FT CONFLICT 153 183 /FTID=VAR_008855.
 FT CONFLICT 270 270 MISSING (IN REF. 2).
 FT CONFLICT 473 473 MISSING (IN REF. 2).
 SQ SEQUENCE 785 AA; 88381 MW; FE3FC796EC1608E CRC64;

Query Match 34.8%; Score 1652; DB 1; Length 785;
 Best Local Similarity 43.0%; Pred. No. 1.4e-114;
 Matches 345; Conservative 128; Mismatches 258; Indels 71; Gaps 17;
 104 LLGGLLHG--SSGSPG--PSVPLRLSYRDLSSANSAIFLGPGSLNLAQMYLDEYR 160
 7 LLMASLLTGAMPSPPTODHLPATPRVRLISFKELKATGTAHFNFLLNTDTYRILKKDDH 66
 161 DRLEFLGIDALYSRLDAQMPDPREVLMPPPOGQRECVKRGRLPILCANFVYLOPHN 220
 67 DRMYVSGNDVYLSLDLDINREPLIIMASPNRIEVCISKDVNGECGNFVRLIOPWN 126
 221 RTHLLACGTGAFOPTCALITYGHRG-----EHW 248
 127 RTHLYVGTGAYNPMCTYYVNGRRAQATPMTQGVAVGRSGRATDGLLRMPATPRDXYI 186
 249 LHLPEVSGVSGRGCRPHEPSRPFSTFIDGELYTLTADFLGREAMIFRSGGPRALRSD 308
 187 FYLEPERLESGKCPYDPKLTASALINELVAGVYIDEMGTDAIFRTLGKOTAMRTD 246
 309 SDOS-LLHDFEFVMAARIPENSDQNDKVFSEFSETPSPDGSNHTVYRGRVVCYND 367
 247 QYNSRWLNDSFTHAELIPDSAEKNDOKLYFFRR--SAAPOSFAVYVARIGRICLND 304
 368 GGRVLYNKKSTLKLKATLVCSVPGCAETHPDQLEDVFLIMPKAKSLEVALFTVSA 427
 305 GGHCCLVKKSTFLKALVCSVPGEDGIEHFDLDFVQOTQDVANPVYAVFTSSGS 364
 428 VFQGFVAVCYVHMADIWEVFGPRANRDPQHONGPGKVPFRPGVCPBKMTAOGRRP 487
 365 VFRGSAVCYVSMADIMVFNFRPAHKEGRYOMPFSGKMPYRPGTCPEG-TFTYS-M 421
 488 GSTRKQDPDEVLOFARAPLMEFVPRHGRVLYKTHLAQQLHQIYVDREAEADTYDYI 547
 422 KSTRKQDPDEVINEMRSHPLMYQAVYPLQRRPLVFRGAPRLITTVADVDADGYEVL 481
 548 FLGTDSGYLVKVALQAGSAPEEYVLELOVEKVPYITEIEISVKQMLYGSRLCY 607
 482 FLGTDRGYQKIVLPK-DQELELELEVEYFKDPAVKVTWTJSSKQQLYVASAVGV 540
 608 AOLRLHOCETVGTACAECLARDPYCAMDGASCTHYRPSLGRFRPRDRIHGNPALQCL 667
 541 TRLSLHRCQAYGAACADCLARDPYCAMDGASCTHYRPSLGRFRPRDRIHGNPALQCL 667
 668 G--OSOEAEVGLVAATMYVGTENHSTFLECLPKSPQAAVRLDOR-PGDEGPDVYKIDE 724
 600 GPNANANKNAV-----ESVQYGVAGSAAFLECPRSPOATVKWLFORDPDR-RREIRAD 654
 725 RVLATRGILFRRLSRFAGVYCTTLEHGSQTVRLALVTVASOLD-NLEPPEKPE 783
 655 RLRTREGILLRLALQDSRGVSCATENNFKHYVTRQVLANHGDVAHNAALFPLLSMA 714
 784 EPPARGGLASTPPKAWYDIQILGFANLPYDEVCEYVWCR--GTTECGSCFRSRSG 840
 715 PPPPGAG---PPLPPYOLAQLAQPEVGLIHOYCOGWHRVPPSPREAPGAPSPBPQ 770
 841 KQARCKSWAGLELGGKMKSRVH 862
 771 DQ-----KKPRNRH 780

Search completed: October 9, 2002, 14:22:39
 Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 13:17:28 : Search time 68 seconds
(without alignments)
2226.038 Million cell updates/sec

Title: US-09-813-290-2
Perfect score: 4746
Sequence: 1 MACALAGKVFPGSWPVMHK.....KKKSRVHADNRTPREVAT 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4215	88.8	782	4	Q9NS98
2	3913	82.4	725	4	Q9H703
3	2052.5	43.2	775	11	Q9QX23
4	1587	33.4	635	4	Q96GX0
5	941.5	19.8	457	4	Q9HBR1
6	897.5	18.9	296	11	Q9J129
7	876.5	18.5	963	4	Q9C0C4
8	860.5	18.1	893	4	Q9C0B8
9	697.5	14.7	762	4	Q9H3S1
10	688	14.5	687	4	Q9BXR8
11	685.5	14.4	1030	4	Q9H2E6
12	685.5	14.4	1049	4	Q9P2H9
13	685	14.4	284	11	Q54948
14	682.5	14.4	1022	4	Q9P249
15	657	13.8	963	11	Q91Y36
16	649	13.7	935	4	Q96JF8

17	640	13.5	1005	11	Q9E071	Q9eq71 mus musculu
18	639.5	13.0	287	4	Q9UDQ1	Q9udq1 homo sapien
19	618	13.5	517	4	Q9NRK9	Q9nrk9 homo sapien
20	610	12.9	724	5	Q9V707	Q9v707 drosophila
21	596.5	12.6	770	5	Q44253	Q44253 drosophila
22	596.5	12.6	770	5	Q9V3M4	Q9v3m4 drosophila
23	578	12.2	418	4	Q96J09	Q96j09 homo sapien
24	574.5	12.1	1202	4	Q9P283	Q9p283 homo sapien
25	561.5	11.8	328	13	Q9YHX3	Q9yhx3 brachydanio
26	541	11.4	367	4	Q9H4H9	Q9h4h9 homo sapien
27	520.5	11.0	616	5	Q9V7P8	Q9v7p8 drosophila
28	441.5	9.3	653	12	Q64906	Q64906 aleoelaphine
29	421.5	8.9	1083	5	Q9V7T0	Q9v7t0 drosophila
30	407.5	8.6	1081	5	Q9U631	Q9u631 drosophila
31	377	7.9	510	4	Q9NX92	Q9nx92 homo sapien
32	370.5	7.8	475	4	Q96FK5	Q96fk5 homo sapien
33	364.5	7.7	612	12	Q9J5F6	Q9j5f6 fow1pox vir
34	332.5	7.0	658	5	Q95XP4	Q95xp4 caenorhabdi
35	323.5	6.8	658	5	Q9N3J8	Q9n3j8 caenorhabdi
36	322.5	6.8	424	4	Q9NS35	Q9ns35 homo sapien
37	321.5	6.8	676	5	Q9TK64	Q9tk64 caenorhabdi
38	305	6.4	699	4	Q96SW4	Q96sw4 homo sapien
39	225	4.7	1841	4	Q15031	Q15031 homo sapien
40	217.5	4.6	1894	11	P70206	P70206 mus musculu
41	209.5	4.4	328	4	Q9HA40	Q9ha40 homo sapien
42	207	4.4	264	5	Q95Q16	Q95q16 caenorhabdi
43	190	4.0	1945	5	Q96681	Q96681 drosophila
44	190	4.0	1945	5	Q9V491	Q9v491 drosophila
45	189.5	4.0	1754	4	Q9UW2	Q9uw2 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9NS98	PRELIMINARY:	PRT:	782 AA.
AC	Q9NS98	01-OCT-2000 (TREMBlrel. 15, Created)		
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DF	01-DEC-2001 (TREMBlrel. 19, Last annotation update)			
DE	SEMAPHORIN SEM2.			
GN	SEM2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,			
RT	Saito T.,			
RL	"Human semaphorin."			
DR	Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL, AB029496; BAA98132.1; -			
DR	InterPro; IPR003599; IG_MHC.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR003659; PSI.			
DR	Pfam; PF00047; Ig_1.			
DR	SMART; SM00409; IG_1.			
DR	SMART; SM00423; PSI_1.			
SO	SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;			

QY	94	MAPSMAATCWLGLGILHGGSSGSPGSPVPRRLSYRDLASNSAIFLPGQSLNQA	153
DB	1	MAPSMAATCWLGLGILHGGSSGSPGSPVPRRLSYRDLASNSAIFLPGQSLNQA	60
QY	154	MLDEYRDLFLFGIDALSLRLDQAMPDPREVLPMPGQREECVKKGRDPLTECANFV	213
DB	61	MLDEYRDLFLFGIDALSLRLDQAMPDPREVLPMPGQREECVKKGRDPLTECANFV	120

```

QY 214 RVLQPHNRTHLLACGTGAFOPTCALITYGHRGEHVLHLEPGSVESGRGRCRHPSPRPAS 273
DB 121 RVLQPHNRTHLLACGTGAFOPTCALITYGHRGEHVLHLEPGSVESGRGRCRHPSPRPAS 180
QY 274 TFIIDELTYGLTADFLGDEAMIFRSGGPRPALRSDSDSLHDPFRVMAARIPENSDDN 333
DB 181 TFIIDELTYGLTADFLGDEAMIFRSGGPRPALRSDSDSLHDPFRVMAARIPENSDDN 240
QY 334 DKVYFEFSETPSPDGSNNHTVSRVGRVCVNDAGGQRYLVNKKSTFLKARLYCSVPBG 393
DB 241 DKVYFEFSETPSPDGSNNHTVSRVGRVCVNDAGGQRYLVNKKSTFLKARLYCSVPBG 300
QY 394 GAETHFDLEDFLLMPKAGSLEYALFSTVSAVFOGFAVCVYMAIMWEFNPFNRH 453
DB 301 GAETHFDLEDFLLMPKAGSLEYALFSTVSAVFOGFAVCVYMAIMWEFNPFNRH 360
QY 454 DGPQHQMGPRYGGKVPFPPRGVCPSPKMTAOPGRPFEGSTKDYDDEVLOQFARHPILMPVPR 513
DB 361 DGPQHQMGPRYGGKVPFPPRGVCPSPKMTAOPGRPFEGSTKDYDDEVLOQFARHPILMPVPR 420
QY 514 RHGRPVLYKTHLAQOLHOIYVDRAEDGTVDYFLGTDSGSYLKVILOAGGSAPEBEV 573
DB 421 RHGRPVLYKTHLAQOLHOIYVDRAEDGTVDYFLGTDSGSYLKVILOAGGSAPEBEV 480
QY 574 VLEELQYFKVPTITEMEISVYKRMVYSGRLGYAQLRLHOCETVGTACACCLARDPYC 633
DB 481 VLEELQYFKVPTITEMEISVYKRMVYSGRLGYAQLRLHOCETVGTACACCLARDPYC 540
QY 634 AMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLIGOSQEEBAVGLVATWVYGTENSTF 693
DB 541 AMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLIGOSQEEBAVGLVATWVYGTENSTF 600
QY 694 LECIPKPSQAAVRLWLRPDEGPDQVKTDERVLTHERGLFRRLSRDAGTYCTTLEH 753
DB 601 LECIPKPSQAAVRLWLRPDEGPDQVKTDERVLTHERGLFRRLSRDAGTYCTTLEH 660
QY 754 GFSQTVVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYKDILQILGFA 813
DB 661 GFSQTVVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYKDILQILGFA 720
QY 814 RVDEYICRWKRCGTTEGSGCRSRSRGKQAGKSMAGLELGKKMSRVHAHNTPRVE 873
DB 721 RVDEYICRWKRCGTTEGSGCRSRSRGKQAGKSMAGLELGKKMSRVHAHNTPRVE 780
QY 874 AT 875
DB 781 AT 782

```

RESULT 2

```

ID 09H703 PRELIMINARY; PRT; 725 AA.
AC 09H703;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
FLJ00014 PROTEIN (FRAGMENT).
GN FLJ00014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA OHara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
spleen."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024425; BAB15715.1;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig-like.

```

```

DR InterPro; IPR003606; Ig_MEC.
DR InterPro; IPR003659; PSI.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00409; Ig_1.
DR SMART; SM00410; Ig-like_1.
DR SMART; SM00423; PSI_1.
FT NON_TER 1
SEQUENCE 725 AA; 80801 MW; BEC8B88FC580D743 CRC64;

```

```

Query Match 82.4%; Score 3913; DB 4; Length 725;
Best Local Similarity 100.0%; Pred. No. 5,5e-317;
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 151 LOAMLYDERDLFLFGIDLALSLRLDQAMPDREVLMPPOGQRECVKGRDPLETCA 210
DB 1 LOAMLYDERDLFLFGIDLALSLRLDQAMPDREVLMPPOGQRECVKGRDPLETCA 60
QY 211 NFVRVLOPHNRTHLLACGTGAFOPTCALITYGHRGEHVLHLEPGSVESGRGRCRHPSPRP 270
DB 61 NFVRVLOPHNRTHLLACGTGAFOPTCALITYGHRGEHVLHLEPGSVESGRGRCRHPSPRP 120
QY 271 FASTFIDELTYGLTADFLGDEAMIFRSGGPRPALRSDSDSLHDPFRVMAARIPENS 330
DB 121 FASTFIDELTYGLTADFLGDEAMIFRSGGPRPALRSDSDSLHDPFRVMAARIPENS 180
QY 331 QDNCKVYTFESETVSPDGSNNHTVSRVGRVCVNDAGGQRYLVNKKSTFLKARLYCSVP 390
DB 181 QDNCKVYTFESETVSPDGSNNHTVSRVGRVCVNDAGGQRYLVNKKSTFLKARLYCSVP 240
QY 391 GPGAEETHFDLEDFLLMPKAGSLEYALFSTVSAVFOGFAVCVYMAIMWEFNPF 450
DB 241 GPGAEETHFDLEDFLLMPKAGSLEYALFSTVSAVFOGFAVCVYMAIMWEFNPF 300
QY 451 AHRDGPQHQMGPRYGGKVPFPPRGVCPSPKMTAOPGRPFEGSTKDYDDEVLOQFARHPILMP 510
DB 301 AHRDGPQHQMGPRYGGKVPFPPRGVCPSPKMTAOPGRPFEGSTKDYDDEVLOQFARHPILMP 360
QY 511 VRPHGRPVLYKTHLAQOLHOIYVDRAEDGTVDYFLGTDSGSYLKVILOAGGSAPE 570
DB 361 VRPHGRPVLYKTHLAQOLHOIYVDRAEDGTVDYFLGTDSGSYLKVILOAGGSAPE 420
QY 571 EEVVLQYFKVPTITEMEISVYKRMVYSGRLGYAQLRLHOCETVGTACACCLARD 630
DB 421 EEVVLQYFKVPTITEMEISVYKRMVYSGRLGYAQLRLHOCETVGTACACCLARD 480
QY 631 PYCAMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLIGOSQEEBAVGLVATWVYGTEN 690
DB 481 PYCAMDASCTHYRPSLGRKRRFRRODIRHGNPALQCLIGOSQEEBAVGLVATWVYGTEN 540
QY 691 STEFELPKPSQAAVRLWLRPDEGPDQVKTDERVLTHERGLFRRLSRDAGTYCTT 750
DB 541 STEFELPKPSQAAVRLWLRPDEGPDQVKTDERVLTHERGLFRRLSRDAGTYCTT 600
QY 751 LEHGFQTVVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYKDILQILGFA 810
DB 601 LEHGFQTVVRLALVIVASQDLNLFPEPEKPEEPARGGLASTPPKAWYKDILQILGFA 660
QY 811 NLPVDEYICRWKRCGTTEGSGCRSRSRGKQAGKSMAGLELGKKMSRVHAHNTPR 870
DB 661 NLPVDEYICRWKRCGTTEGSGCRSRSRGKQAGKSMAGLELGKKMSRVHAHNTPR 720
QY 871 EPEAT 875
DB 721 EPEAT 725

```

RESULT 3

```

ID 09QX23 PRELIMINARY; PRT; 775 AA.
AC 09QX23;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)

```

DE SEMAPHORIN M-SEMAK.
 GN SEMA3E.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK 6.
 RA Miyazaki N., Furuyama T., Inagaki S.;
 RT "A novel semaphorin, M-Semak which inhibits neural outgrowth from
 RT sensory neurons."
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF034744; AD01996.1; -
 DR MGD: MGI:1340034; Sema3e.
 DR Interpro: IPR003599; 1g.
 DR Interpro: IPR003006; 1g_MHC.
 DR Interpro: IPR003659; PSI.
 DR Interpro: IPR001627; Sema.
 DR Pfam: PF00047; 1g. 1.
 DR Pfam: PF01403; Sema. 1.
 DR SMART: SM00409; IG. 1.
 DR SMART: SM00423; PSI. 1.
 DR SMART: SM00423; PSI. 1.
 SQ SEQUENCE 775 AA; 89543 MW; 221E766F404098D4 CRC64;

Query Match 43.2%; Score 2052.5; DB 11; Length 775;
 Best Local Similarity 50.7%; Pred. No. 4,6e-162;
 Matches 399; Conservative 127; Mismatches 234; Indels 27; Gaps 10;

OY 94 MARSAMALCWLGLLHGGSSGSPGSPVRLRLSYRDLISANSALFPGSLNQA 153
 DB 1 MAPAGHITLTLMLLHLEMTPLGHSANPSYRLRLSHKLELNTSTJFOSPLGLDHT 60
 OY 154 MYLDEYRDLRLGLGADALYSRLDQAMPDPREVLMPPOGOREECVRKGRDPLTECANFV 213
 DB 61 MLDBYORLRLVGGGRDLVSLNLEVSQYREITWPSYAVKVEECIMKGD-ANCCANYI 119
 OY 214 RVLORPHNTHLLACGTGAFQPTCALITVGHGHEVL-HLEPGSVESGRCRCHPERSRPA 272
 DB 120 RVLHNYNTHLLTCAFGAFPHCAIRVGHHSERPLFHLESRSRGRCRCPFDENSFV 179
 OY 273 STFDIGELTYGLTADFLCREAMIFRSGRPRLRSD-SDQSLHDPFRYMAARITENSQ 331
 DB 180 STLVNELLFAGIYSDYWRGDSAFIRSMGKLGIRTEHDERLKEPKFVSGIMRIDNEPR 239
 OY 332 DNDKYEFSEFVPSPDGGSNHTVSRGVYCVNDAGGGRVLYNKMSTFLKARLYCSVPG 391
 DB 240 DNNKTYFFTERALFAENNA-HTTYRGRGLCVNDMGGRILVYKWSFTLKARLYCSVPG 298
 OY 392 PGGAETHFDQLEDFVLLMPKAGKSLEYVALFSTVSATVPOGFANVCYVHMADIWEVNGPRA 451
 DB 299 MNGIDTYDELEDFVLLRDPKNPVIGLFNTSINIFRGHNAVYVHMSSIREAFNGPYA 358
 OY 452 HRDGRQHOHMGPRGVPRPSPGSCSKMTAOPRGPGSKTQDPDEVLOARAHPLMFV 511
 DB 359 HKEGGEYHMSLVEGKVPYPRPGSCSKVN--GKYYGTTKPPDAIRFARHNPMLYOP 415
 OY 512 RRRHRPVLVKTHLAQOHLQIVDRAVEDGTVDVIFLTDGSGVLYKVALOAGSABE 571
 DB 416 KVVHKKPILVKTGDKYNLRQLAVDRAVEDGQYDLFITDGIYALKVITITINOETEMNE 475
 OY 572 EYVLEELQVFKVPTITEMEISVKRQMLVYGSRLGVALRLHQCETTYGTACAECCCLARP 631
 DB 476 EYVLEELQVFKVPTITEMEISVKRQMLVYGSRLGVALRLHQCETTYGTACAECCCLARP 631
 OY 632 YCAMOGASCTHNRP--SLGKRPRRRODINHGPNALQCLGOSOEBAVGLVATVNYGTEH 689
 DB 536 YCAMOGISCSRYTPGANAHRRRRRRVDVHGNAAQCFQOQVGDVLDLRTERLAYGIS 595
 OY 690 NSTFLECLPKSPQAAVRMLLQRPDGGDPQVYTDERVLTERGLRRLSRDAGTYTCT 749
 DB 596 NSTLLECLPKSRQAKVIWVQKGRDVRKEGVATDDRVKMDGLLFLVRKKSAGTYTCT 655

OY 750 TLEHFSQTVLRALVIVASQIDNLPPEKPE-----PARGLASTPPKAMTKDI 803
 DB 656 TVEHNFVHTVAKITLLEVEEHEKHEMKHDEEERHNMPCPPISGMSQGTKP--WYKEF 713
 OY 804 LQIGFANPLPVDEYCEVWRCRTTEGSGCRSRSGOAGKSMAGLELKKKSRVHA 863
 DB 714 LQIGSNFQVREYCEKWCSTD-----KKRKLKMSPKMYANPOEKRLSK--A 763
 OY 864 EHNTRPR 870
 DB 764 EHNTRPR 770

RESULT 4
 ID 096GXO PRELIMINARY; PRT; 635 AA.
 AC 096GXO;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE UNKNOWN (PROTEIN FOR MGC:18122)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRIN, AND ANAPLASTIC OLIGODENDROGLIOMA WITH 1p/19q LOSS;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC009113; AA09113.1; -
 DR EMBL: BC009113; AA09113.1; -
 SQ SEQUENCE 635 AA; 70694 MW; D994099B476B9210 CRC64;

Query Match 33.4%; Score 1587; DB 4; Length 635;
 Best Local Similarity 49.9%; Pred. No. 2e-123;
 Matches 321; Conservative 90; Mismatches 196; Indels 36; Gaps 12;

OY 211 NEFVRLORPHNTHLLACGTGAFQPTCALITVGHGHEVL-HLEPGSVESGRCRCHPERSR 269
 DB 2 NEFVRLORPHNTHLLACGTGAFQPTCALITVGHGHEVL-HLEPGSVESGRCRCHPERSR 269
 OY 270 PFASTFIDGELTYGLTADFLCREAMIFRSGRPRLRSD-SDQSLHDPFRYMAARITENSQ 328
 DB 62 RAASVIVGELTSGVAAADMGKDTIFRSLGQRPSTLTPHRSKRLNEKFKVGMIPES 121
 OY 329 SPQDNDKYEFSEFVPSPDGGSNHTVSRGVYCVNDAGGGRVLYNKMSTFLKARLYCS 388
 DB 122 ENPDDKTYFFTERALFAENNA-HTTYRGRGLCVNDMGGRILVYKWSFTLKARLYCS 181
 OY 389 VPFGAETHFDQLEDFVLLMPKAGKSLEYVALFSTVSATVPOGFANVCYVHMADIWEVNG 448
 DB 182 VPGVEG-DTHFDQLEDFVLLSRSRDNHTPLLYAVFT--SIFGGSACVYVSMNDVRAFLG 239
 OY 449 PRAHNRQHOHMGPRGVPRPSPGSCSKMTAOPRGPGSKTQDPDEVLOARAHPLMFV 508
 DB 240 PRAHNRQHOHMGPRGVPRPSPGSCSKMTAOPRGPGSKTQDPDEVLOARAHPLMFV 508
 OY 509 WYVRPHRGPVLVKTHLAQOHLQIVDRAVEDGTVDVIFLTDGSGVLYKVALOAGSABE 568
 DB 296 NSVLPGRGRLFLQVGNANTFTQIADRVAAADGHVDFLIGDVGTVLKIVSVPGSRP 355
 OY 569 EPEVLEELQVFKVPTITEMEISVKRQMLVYGSRLGVALRLHQCETTYGTACAECCCLARP 628
 DB 356 SAEGILLEELHVEEDSAVTSQISKRQOLVYASAAQILHCAHAGRYCTECCIA 415
 OY 629 RDPYCAMDASCTHNRP--SLGKRPRRRODINHGPNALQCLGOSOEBAVGLVATVNYGTEH 688
 DB 416 RDPYCAMDAGVACIRFGPS-AKRRFRQDVANRDPSTLCSGDSR--ALIEHKVGEV 470
 OY 689 HNSTFLECLPKSPQAAVRMLLQRPDGGDPQVYTDERVLTERGLRRLSRDAGTYTCT 748
 DB 471 GSSAFLECEPERSIOARVETFRAGVTAHVOVLAETERTARGLLRLRRDSSGYLC 530

DE	SEMAPHORIN (FRAGMENT).
OC	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	(1)
RP	SEQUENCE FROM N.A.
RC	STRAIN-SPRAGUE-DAWLEY;
RA	Breuz Verica M.S., Widmer D.A.J., Wagner G.C., Dreyer J.L.:
RT	"New rat semaphorin with closest homology to avian Sema3D (C-
RL	collapsin-2)".
DR	Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
FT	NON_TER
FT	NON_TER
SQ	SEQUENCE 296 AA; C2226212885DD044C CRC64;
Query Match	18.9%; Score 897.5; DB 11; Length 296;
Best Local Similarity	57.0%; Pred. No. 1.7e-66;
Matches 174; Conservative	51; Mismatches 67; Indels 13; Gaps
QY	332 DNDKYFFEFSEVPEPDDG-SNHVVNVSRGRGVNDAGSGORLVNKMSTFLKARLYCVP 390
Db	1 DDDKTIFFFPRES--SQEGTSDRSILSRGRCKMDVGQRSLINKMTWTFKLARLVCSIP 58
QY	391 GPGAEHTHEDOLEDFVLMLPKAKGLEEVATLEPTVSAAVEGFAVCVYNHADIMEVFNGCF 450
Db	59 GSDGADTHFDELQDIYLPLPTRBERNPVYGVFTTSSIFKSAGCAVCYSTADIARAENGY 118
QY	451 AHRDGFOMQMPGGKVPPRRPGVCPSKMTAOPGRFGSGTKDYPRDLQAFANPLMFWP 510
Db	119 AHKEADNRWQYDGRIEYPARGTCPSK-TYPD-LIKSTRFPDPDVISFIIRHFVMFKS 175
QY	511 VPRGRGRVLVKTHLAOOLHQIVLDRVVAEDGTVDYVIFLTGDSGSVLKVIALQAGSAEP 570
Db	176 YVPVAGAPTPQLINDYRLQLQIVDHVAABEQQIYVMELGTDGXVLYAVSI---SKEK 231
QY	571 ---ENVYLEELQVKKVPIPTIMETLSYKRQMLYVSGSLGVALRLHGCEYTGACAECL 627
Db	232 WNMEEVLEELOVEKHPPAILMNEELSLKQOOOLYVGSMDGLVOLSHRDYTGKACADCCL 291
QY	628 ARDPY 632
Db	292 ARDPY 296
RESULT 7	
ID Q9C0C4	PRELIMINARY; PRT; 963 AA.
AC Q9C0C4;	
DT 01-JUN-2001 (TREMBLrel. 17, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE KIAA1739 PROTEIN (FRAGMENT).	
GN KIAA1739.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN (1)	
RP SEQUENCE FROM N.A.	
RX MEDLINE=21082932; Pubmed=11214970;	
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;	
RT "Prediction of the coding sequences of unidentified human genes. XIX.	
RT The complete sequences of 100 new cDNA clones from brain which code	
RT for large proteins in vitro."	
RL DNA Res. 7:347-355(2000)."	
EMBL: AB051526; BAB21830.1; -.	
InterPro: IPR003599; Ig.	
InterPro: IPR003600; Ig_Like.	
InterPro: IPR002165; Plexin_repeat.	
InterPro: IPR003659; PSI.	
Pfam: PF01437; Plexin_repeat; 1.	

Db 602 SSSCKHVSIXPOLATRPW-IODIEGASAKDLCASSVSPSPVTEGKCEQYQFQPN 660
 QY NSTFLECPKSPQAAVRMLQRPDEGPOVKTDERVLHTEGGLFRRLSRDPACTYTC 749
 Db 661 TWTNLTACCLSLATRLML--RNG--APVNASASCHVLTGDL--VGTOOLEFQCM 713
 QY 750 TLEHGSQTIVRLALVT--VASOLD 773
 Db 714 SLEEGFOOLVASYPEVEDGVADQTD 740

RESULT 9

Q9H3S1 PRELIMINARY; PRT; 762 AA.
 ID Q9H3S1
 AC Q9H3S1
 DT 01-MAR-2001 (Tremblrel, 16, Created)
 DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
 DE SEMB.
 GN SEMB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.,
 RT "Human semaphorin B."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB029394; BAR20087.1;
 DR InterPro: IPR000005; HTHATAC.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR003659; PSI.
 DR Pfam: PF01437; Plexin_repeat. 1.
 DR PRINTS: PR00142; RECA.
 DR SMART: SM00423; PSI. 1.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; UNKNOWN.1.
 DR SEQUENCE 762 AA; 83195 MW; 6A738E4EBE834D4 CRC64;
 SQ

Query Match 14.7%; Score 697.5; DB 4; Length 762;
 Best Local Similarity 28.4%; Pred. No. 2.8e-49;
 Matches 236; Conservative 125; Mismatches 326; Indels 145; Gaps 35;

QY 92 PSAPASMAIICWLLGILLH-----GSSGSPSPVPRRLSYRDLNANSAIF 142
 Db 4 PALGLDPWSILGLFLFOLLPTTATAGGGGCGP-----MPRYRYAGD---ERRALSF 55
 QY 143 LGPGSLNQAMYLDEYRDLFLGGLDALYSLY-DOAMPDPREVLPPOPGORECVR 200
 Db 56 FHOKGLQDEDTLLLSGDGNTLVYGARALIALDIOPGVPRLNMIWPASDRKSKSCAF 115
 QY 201 KGRDPLECANFVRLVLOPHNRTHLACGAFQPTCALIVHGRGHEVHLLEGVSSEGR 260
 Db 116 KKSNETQCFNFRVLVSVVTHLYCGTFAFPACTFTEL--QDSTYLDISEDKWMEG 173
 QY 261 GCRPHRSPRPASTFDGLYTGTLADFLGREAMIFRSGGPRALRSDD--QSLHDPFR 319
 Db 174 GOSPEPAPKHTAVLDGMLYSGTNNFLGSEPIILMRILGSCVLTLDNFRMLHNDASF 233
 QY 320 VAAARPEPSDDNDNVYFEFSTVSPSPGGSNHYVSVRGRCVADAGQARLVNKMST 379
 Db 234 V--AATP-----STGVYVYFEFETASEPD-FERLHTSRVARKCKNDVGKELLOKKWT 285
 QY 380 FLKARLVCSVPPGGAETH-----FDLEDVFLIMPKAGKSLLEVYALFTVSAY 428
 Db 286 FLKADLT-SAPRGSCSPSSSATRSCSPILRLQLPTSTQSSPSSGQ-----VGTGRSS- 337
 QY 429 PGGFVAVYHMDIMEVFNPRFAHNRGPHOGKVPYGGKVPFRPGVCSPKMTAQPRPFG 488
 Db 338 ---AVCAFLSLDIEVERFGKREKLNKETSRTYTYRGPETNRPSPGSC----- 380

QY 489 STKDYIDEVLOFAAHPLIMFWPRHGRVLYKTHLAQOLHOIYDVRYEADG--TYDYI 547
 Db 381 SVGPSDDKALFTFMKDHFLDEGV---VGPRLVYKSGV--EYRLAVETAQGLDGHSLWM 435
 QY 548 FGTDSGSVLYKVIYALQAGSAPDEEVLELOVFNVPRTIMEISYKQMLYVSRLCV 607
 Db 436 YIGTTTSGLSHKRAV-----SGDSSAHVLEIOFLPDEVRNLQALPQGAIVFVPSGV 490
 QY 608 AOLRLHOCETVGTACAECCLPARDPYCAMDGAS---CTHYRPSLGRKRRFRDRIKGNAL 664
 Db 491 WRVPRANCNVY-ESCVDCVLPARDPCAMPDESRQCCLLSAPRLNS---WKQDMENGENW 546
 QY 665 QCLGSGQEEANGLVAAIM-----YGTGHNSTFLECLPKSPQAAVRMLQRP 713
 Db 547 AC-----ASGPMRSRLRQSRPQIIKEVLAIVNSLILEPCPHLSALASYW----- 592
 QY 714 DEGPQVKTDERVLTHTGGLFRRLSRFDAGYTCGTTLEHGSQVYVRLALVVIYASOLD 773
 Db 593 SHGPAAYPEASTVY--NGSILLIYQDGVGLYQCWATENGSTYVIS---YVWDSQ-D 645
 QY 774 NLFPPEPKPEEPAPAGLSTPPKAMYKDILQIGFANLPRVDEXCEBWCNGT----- 828
 Db 646 QTLADPE-----LAGIPREHVKVLTRVSGGALAAQGSY---WHFVTVTLF 692
 QY 829 --ECGCGF-----RSRKRQKARGK--SWAGLELGKKMK-SRYHAENHRTPRE 871
 Db 693 ALVYSGALILIVASPLRALRARGKVGCGETLRPGEKAPLSR--EQLHOSPRE 742

RESULT 10

Q9BX8 PRELIMINARY; PRT; 687 AA.
 ID Q9BX8
 AC Q9BX8
 DT 01-JUN-2001 (Tremblrel, 17, Created)
 DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
 DE SEMAPHORIN 6B ISOFORM 2.
 GN SEMA6B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21248680; PubMed=11350127;
 RA Correa R.G., Sasahara R.M., Bengtson M.H., Katayama M.L.H.,
 RA Salim A.C.M., Brentani M.M., Sogayar M.C., de Souza S.J.,
 RA Simpson A.J.G.,
 RT "Human Semaphorin 6B (HSA) SEMA6B", A Novel Human Class 6 Semaphorin
 RT Gene: Alternative Splicing and All-Trans-Retinoic Acid-Dependent
 RT Downregulation in Glioblastoma Cell Lines."
 RL Genomics 73:343-348(2001).
 DR EMBL: AF293363; AK16831.1;
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR00737; Squash.
 DR SMART: SM00423; PSI. 1.
 DR SMART: SM00286; PTL. 1.
 DR SEQUENCE 687 AA; 74967 MW; CBC68CF5D69738D7 CRC64;
 SQ

Query Match 14.5%; Score 688; DB 4; Length 687;
 Best Local Similarity 32.0%; Pred. No. 1.5e-48;
 Matches 203; Conservative 91; Mismatches 262; Indels 78; Gaps 24;

QY 90 QPSPMAIICWLLGILLHGGSSGSPSPVPRRLSYRDLNANSAITLG----- 144
 Db 2 QTPRAPPPRALLL--LTLGGAHGLFP-EPPPLSVAPROYL--NHYPVVGSGPGR 55
 QY 145 ---PQSLNQAMYLDEYRDLFLGGLDALYSLRLQAMPD-----PREVLMPPOPGORE 197
 Db 56 LTPAGADLDINQRLVNRVLTLEIGDRDNLRYVELEPPTSTELRYOKLTWNSPESDIV 115
 QY 198 CYRKGDPLECANFVRLVOPHNKTHLACGAFQPTCALITVHGRGHEVHLLEP-GSV 256

```

Db 116 CRMKGKE-GEGRNFVKYLLRDESTLFVCGSMNFPCANYSID-----TLQPGDN 167
QY 257 EGRGRCHPEPRPASTFICGELTGLADFLGREAMIFRSGRPRLARS-DSQSLIH 315
Db 168 ISGMARCPDRPKHANVALESDMLTATVTDLADAVYISLGRPTLRKYKHSKMF 227
QY 316 DREFVMAARIPENSDDNDKYFFFESETVPSPDGSNHYTVSRGVCVNDAGO-RVLY 374
Db 228 EYFYVAIV-----EMSHVYEFERE-IAMFNYLEKVVSRVAVCKNDVGSGSPYLE 279
QY 375 NKWSTFLKRLVCSVPGGGAETHEDOLEDFLLMPKAGKSLLEYALFSTYSAVPGQAV 434
Db 280 KQWSTFLKRLKCSV--PQDSHFYFENVLDAYGVVSLGGRV-VLAIVSTPSNLSGSAV 336
QY 435 CVYHAMDIYEVNFGPAHNDGPOHONGPY-GGKVPFRPGVCSKMTAOPGRPGSTKDY 493
Db 337 CAFDLIQAIVAVFGRFRREKSPESITVPRBEDQVPRPRGCC-----AARGMAYNASSAL 391
QY 494 PDEVLOFARAHPLMPVPRPRHGRVLYKTHLAQQLHQTVDVREAEEDGTVDYIFLGTS 553
Db 392 PDLINFEVKTPIHMDAIVPSLGHAPWILRTLMHQLTRAVADVAGAPWGNQTVFLGSEA 451
QY 554 GSVLKVI-----ALQAGSAPPEEVLEELQVFK-----VPPITEMEIS 593
Db 452 GTVLKFLVRPNASTSGTSG--LSVPLEEFETVPRDRCERPGGETGTQRLDLSLEID 509
QY 594 VKROMLYVGRSLVADRLHCEETGTACAECLLADPYCAW--DGASCTHYRPSLGKR 651
Db 510 AASGGLAARPCVVRVAVRARCQOYSGCMKNCIGSDPYCGMAPDG-SCIFLSP--GTRA 566
QY 652 FRRODIRHGNPA--LQCLG---QSOEEAVGLYA 680
Db 567 AFEQDVASGASTSGIGDCTGLLRASLSERAGIVS 600

```

RESULT 11

```

ID 09H2E6 PRELIMINARY; PRT; 1030 AA.
AC 09H2E6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SEMAPHORIN SEMA6A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564339; PubMed=10993894;
RA Klostermann A., Lutz B., Gerlier F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/SEMA6A-1) bind to the Enabled/vasodilator-stimulated Phosphoprotein-1-like Protein (EVL) via a novel carboxyl-terminal Zyxin-like domain."
RT J. Biol. Chem. 275:39647-39653(2000).
RL EMBL: AF279656; AAC29378.1; -.
DR InterPro: IPR003659; PSI.
DR SMART: SM00423; PSI.
SQ SEQUENCE 1030 AA; 114368 MW; A57B79C10AEC4B34 CRC64;

```

Query Match 14.5%; Score 686; DB 4; Length 1030;
Best Local Similarity 30.3%; Pred. No. 3.9e-48;

Matches 199; Conservative 93; Mismatches 278; Indels 86; Gaps 24;

```

QY 104 ILGGLLHGSSGSPGSPVRLRLSYRDLISANRSAIFLPGQSLN-----LQAMYLD 157
Db 7 LLYFTLLHFAAGG--PPEDEP--ISISHGNV--TKQYVAVGKHKPGNTTQRRHRLDIQIM 62
QY 158 EYDRFLFLGGLDALYSRLDQAMPD-----PREVLMPPGRGRECVRKGRDPLETCANFV 213
Db 63 IMMGTLYIARHDHIVYDIDTSHTEELYSCKLTKWKSROADVDYTCMKGKHK--DECHNFI 121
QY 214 RVLDPRHRTLLACGTGAPORTCALITVGRGHVHLDEP--GSVEGRGRCHPEPRAPA 272

```

```

Db 122 KVLKKMDDDLFCQGTNAFNPSC-----RNYKMDLLEBFGEFEGMARCPYADKANAV 174
QY 273 STFDIGELYGLADFLGREAMIFRSGRPRLARS-DSQSLIHDRFVMAARIPENSQ 331
Db 175 ALFADGKLSATYVDFDLADAVYISLGSPTLRKYKHSKMLKEFYQAV-----D 227
QY 332 DNDVYFFFESETPVSPDGSNHYTVSRGVCVNDAGC--ORVLNWKSTFLKRLVCSVP 390
Db 228 YGDIYTFEERE-IAVENITMGKVPFRVAVQCKNDMGSGQRLKEQWSTFLKRLKCSVP 286
QY 391 GPGGAETH-----DQLEDVFLMPKAGKSLLEYALFSTYSAVPGQAVYHAMDIYEV 445
Db 287 G-----DSHFYFENVLDAYGVVSLGGRV-VLAIVSTPSNLSGSAVATMDIASV 338
QY 446 FNGRPAHNDGPOHONGPY-GGKVPFRPGVCSKMTAOPGRPGSTKDYDEVLOFARAH 504
Db 339 FTGRKREKSPDSTVTPVPRDERVPRKPRGCCAGSSSILE---RVATSNMPPDDTLNFIKTH 395
QY 505 PLMPVPRPRHGRVLYKTHLAQQLHQTVDVREAEEDGTVDYIFLGTSGLVAVLQA 564
Db 396 PLMDAVPSIFRPRWFLRTVRYRLTKIAVDTAAGPYQNHVFLGSEKGLILFLRIG 455
QY 565 GGSAPPEEVLEELQVFK-----VPPITEMEISVKROMLYVGRSLVADRLHCE 615
Db 456 NSGFLNDSLLEEMSVYNSKCSYDGEDKRLMGQDLRASSSLYVAFSTVIVPIGRG 515
QY 616 EYGTACAECLLADPYCAW--DGASCTHYRPSLGKRFRRODIRHGNPAIQLGQSOE 673
Db 516 ERHGCKKTKCIASRPYCGMKEGAGSHLSPN--SRLEFDODIERGN-----T 562
QY 674 EAVGLVATVYGTENSTFLECLERKSPQA-----VRM--LLQRPDESP 717
Db 563 DGIIGDCHNSFVALNGHSSLSLPSTTSDTAQEGYESRGKMLMKHLDDSPSTDP 618

```

RESULT 12

```

ID 09P2H9 PRELIMINARY; PRT; 1049 AA.
AC 09P2H9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE KIAA1368 PROTEIN (FRAGMENT).
DE KIAA1368.
GN KIAA1368.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RC MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro."
RT DNA Res. 7:65-73(2000).
RL EMBL: AB037789; BA892606.1; -.
DR InterPro: IPR003659; PSI.
DR SMART: SM00423; PSI.
FT NON_TER
SQ SEQUENCE 1049 AA; 116511 MW; 7781D20ACC7A8AEA CRC64;

```

Query Match 14.4%; Score 685.5; DB 4; Length 1049;

Best Local Similarity 30.3%; Pred. No. 4.4e-48;
Matches 202; Conservative 94; Mismatches 280; Indels 91; Gaps 26;

```

QY 104 ILGGLLHGSSGSPGSPVRLRLSYRDLISANRSAIFLPGQSLN-----LQAMYLD 157
Db 9 LLYFTLLHFAAGG--PPEDEP--ISISHGNV--TKQYVAVGKHKPGNTTQRRHRLDIQIM 64
QY 158 EYDRFLFLGGLDALYSRLDQAMPD-----PREVLMPPGRGRECVRKGRDPLETCANFV 213

```

```
Query Match      14.4%; Score 685; DB 11; Length 284;
Best Local Similarity 46.2%; Pred. No. 7.0e-99;
Matches 132; Conservative 58; Mismatches 92; Indels 4; Gaps
```

163 LFLGSLDALYSRLDQAMPDRPEVLWMPPOPCRECECRKGGDLPTTECANRVLPDHNT*222
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
1 MYVSGKDVALSLDDHDIRPLLIHMASPORFECTISCKNGRCFGCPNPIIYNYVTCO

[illegible]

Db 325 GVTTLNLSIPGSAVCAFSMDIDIEKVEFKREKQKTPDSVTFVAVPEDKVPRKPPGCCAKH 384
 QY 479 MTAOGRPGSGSKVDVDFVLOFARAHPLMFVPRRGRPVLTHTLAQOLHQIVNDRVE 538
 Db 385 GLAE--AKTSIDPDELSTFIKSHPLMSADAVPIADEPFTKTRVYRLTALISVDHSA 441
 QY 539 AEDGYTVIIFLGTDGSLVLYALOAQGSAPREVEYVELOV-----KVPTEITE 589
 Db 442 GRYQVTVTVFVGEAGMVLKYLAKTSPSLN-DSVLEIEIAYINAKCSAENEDKKYIS 500
 QY 590 MTSIVKRWMLYVGSRLGAOLRLHOCCTEYGTACAECCCLARDPYCAM-DGASCTHYRPSL 647
 Db 501 LQDLKDHNLVYAFSSCITRIPLSKERTGSCCKSCIASRDPYCGMLSGSGGRVTPGM 559

RESULT 15

Q91Y36 PRELIMINARY: PRT: 963 AA.
 ID Q91Y36;
 AC Q91Y36;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SEMAPHORIN 6C.
 GN SEMA6C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C56; TISSUE=BRAIN;
 RA Ou X., Zhai Y., Wei H., Yu Y., Zhang C., Xing G., Zhou K., Zhu Y.,
 RA Ouyang S., He F.;
 RT "Mus musculus mRNA for semaphorin 6C, complete cds."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF63973; AAK52051.1;
 SQ SEQUENCE 963 AA: 102708 MW: 08FA0A3FA6532100 CRC64;

Query Match 13.8%; Score 657; DB 11; Length 963;
 Best Local Similarity 29.1%; Pred. No. 9.2e-46;
 Matches 231; Conservative 89; Mismatches 295; Indels 178; Gaps 33;

QY 95 APSAMAIICWLLGGILHGGSSGSPSPVPR--LRISYRDLISANRSATFLPGQSLINQ 152
 Db 4 APHSMPILLIL--LL--SSLPQAQAFPRDPTPLITSDLGASPSMERGLEDDAVAA 58
 QY 153 AMYLDYR-----DRFLGGLDALYSLRLDQAMPD-----PREVL-WPQPGQRECEYRK 201
 Db 59 ELGLDPQRFLINRTLLVAARDHVESFDL-QAQEGEGLVNKFLLTWRSQ--DMENCAYR 115
 QY 202 GRDPLT-ECANFVAVLOPHNRTHLLAGTGAFOPTC--ALITYGHRGHEVHLHEPGSVE 257
 Db 116 GK--LTDECYNYIVLVWMSQTLACGINSFSPMCRSYGITSLOQEGEEL----- 164
 QY 258 SGRGRCPREPRPASTFIDELTYGLTADFLGREAMIFRSGGRPALRSDS-DQSLIHD 316
 Db 165 SGQARCPDPDATQSTVAIFABESLXSATAADQASDAVYRSLGPPRLRSKAYDSKWLRE 224
 QY 317 PRFYAARIPENSDDNDKYFFPSETPSPIDGGSNHVTVSRGRCVYNDAQO-RVLVN 375
 Db 225 PHEVYAL-----EHGEHYVFFERE-VSVEDARLGRVOFSRVARVCKRDMGSPRALDR 276
 QY 376 KWSITLAKRIYCSVPGGAGTHTFDOLEDDVFLMPKAGKSLEYVALFSTVSAVFOGFAVC 435
 Db 277 HWTSLKRLKNCV--PEDSTTFEDVLSLTGPNVNLHGRS-ALFGVFTTQINSIPGSAVC 333
 QY 436 VYHMADIWEVFPFANDGQOHONGPYG-GKVPFPRPGVCPSKMTAQGRPGSTKDY 494
 Db 334 AYLDLDIENGEGFKEDRSLDGAWTPESEDKVPSRPGSCAGVAA--ASFSSQDLP 390
 QY 495 DEVLOFARAHPLMFVPRRGRPVLTHTLAQOLHQIVNDRVEAEDGTVDYVIFLGTDG 554
 Db 391 DVLLEFIKAPHLDDPAVPATHQPIITLTLSRA-LITVAVDVMAGPHRNTTVLFLGSDG 449

QY 555 SVLKVIALOAGSAPREVEYVELOV-----KVPTEITEISVROMLYGS 603
 Db 450 TVLKVLIP-PEGOSLSGEPVLEEDIDAYSHARCSCGKSPRAARITIELDTEGHRLFVAF 508
 QY 604 RLGAOLRLHOCCTEYGTACAECCCLAR-DPYCAN----- 635
 Db 509 PGCLIVYLSLRCARHG-ACQRCIASLDPYCGWHSRGCMSIRPGGTVDLTGNOESTE 567
 QY 636 -----DGASCTHYRPS-----LGRFRFR-----QDIRGNPALQGLGOSOE 672
 Db 568 HGDCQDATGSQSGPDSAYVLLGPGRSPTPSSPSDAQOG-POSTLGAHTGVRDLS 626
 QY 673 -----EAVGLVAAITMYGTETENSTFLDCLPKSPQAAVAVMLLQRPDGEPOVKT 722
 Db 627 PASASRSIPIPIILLACVAAAFALGASVSGLLVSCACRRANRRRSKDIETPG----- 677
 QY 723 DERVLHTEGILFRRLSRPDAGTYTCTLHEGHSQTVVRLALVYIVASQDNLFPPEPK 782
 Db 678 -----LPRPLSLRLARLHGG-----GREPP 699
 QY 783 EEPFARGGLASTP 795
 Db 700 --PRKDGDAQTP 710

Search completed: October 9, 2002, 14:24:10
 Job time : 75 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 9, 2002, 19:19:59 ; Search time 1640 Seconds
(without alignments)
7201.119 Million cell updates/sec

Title: US-09-813-290-2
Perfect score: 4746
Sequence: 1 MACLAGKVPFMPGSMVPHK.....KMSRVHAHNRPPEVAT 875

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV=xlh
-O/cpnt2.1/USPTO.spool/US09813290/rnat_09102002_094646_19135/app.query.fasta.1.1031
-DB-EST -CPMT-fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09813290_@CGN.1.1180_@rnat_09102002_094646_19135 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estdb:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1625.5	34.2	3436	11	AK014333	Mus muscu
2	820	17.3	647	10	BI067436	pf1n.pko

3	778	16.4	420	9	AM803853	AM803853 IL2-UMC08
4	777.5	16.4	990	10	BI652258	BI652258 603299702
5	749.5	15.8	785	10	BI689389	BI689389 603315834
6	710.5	15.0	645	10	BI066459	BI066459 pf1n.pko
7	701	14.8	662	10	BI57438	BI57438 603238962
8	692.5	14.6	754	9	AL584625	AL584625
9	690	14.5	576	10	BI067451	BI067451 pf1n.pko
10	685.5	14.4	823	10	BI645606	BI645606 603275104
11	684.5	14.4	1095	10	BM473502	BM473502 AGENCOURT
12	665	14.0	1095	9	AL584324	AL584324
13	659.5	13.9	737	10	BI646227	BI646227 603276957
14	658	13.9	609	10	BI066327	BI066327 pf1n.pko
15	652.5	13.7	778	10	BI684444	BI684444 RST3370 A
16	652.5	13.7	920	10	AL541401	AL541401
17	651	13.7	797	10	BI100405	BI100405 602858585
18	641.5	13.5	436	9	BM287610	BM287610 528378 MA
19	637	13.4	748	9	AU136052	AU136052
20	633	13.3	364	9	AU1750874	AU1750874 cno6b10.x
21	629	13.3	674	9	AU138132	AU138132
22	626	13.2	735	9	AU140116	AU140116
23	623.5	13.1	626	9	AU137591	AU137591
24	613	12.9	1121	10	BM477901	BM477901 AGENCOURT
25	612.5	12.9	723	10	BI870437	BI870437 603395690
26	607.5	12.8	774	10	BG215143	BG215143 RST34801
27	602.5	12.7	721	9	AM742133	AM742133 up52f06.Y
28	602	12.7	678	10	BI151058	BI151058
29	600.5	12.7	678	10	AM328571	AM328571 ds03d12.x
30	594.5	12.5	624	9	AU134748	AU134748
31	591.5	12.5	690	9	AU137554	AU137554
32	591.5	12.5	728	9	AU137554	AU137554
33	589	12.4	755	9	AU137539	AU137539
34	586	12.3	637	9	AU135310	AU135310
35	582	12.3	602	10	BM425751	BM425751 pf1n.pko
36	575.5	12.1	918	10	BG252302	BG252302 602365185
37	552.5	11.6	629	9	BI116052	BI116052
38	539.5	11.4	604	10	BI183160	BI183160 UNP-P-FN
39	534.5	11.3	957	10	BG261855	BG261855 602373655
40	530	11.2	933	10	BI686933	BI686933 603313808
41	529.5	11.2	561	9	AT733876	AT733876 z083d09.Y
42	523	11.0	971	10	BG342228	BG342228 602374305
43	518.5	10.9	546	9	AT733905	AT733905 z083h09.Y
44	507	10.7	727	10	BI461517	BI461517 603206395
45	503.5	10.6	701	10	BB615357	BB615357 601280765

ALIGNMENTS

RESULT 1	AK014333	3436 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK014333				
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:3230402E04; sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B, full insert sequence.				
ACCESSION	AK014333				
VERSION	AK014333.1	GI:12852112			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain: C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone: lib:RIKEN full-length enriched mouse cDNA library clone:3230402E04.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Garnier, P., and Hayashizaki, Y.				
TITLE	High efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2 (sites)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to				


```

|||||
Db 684 GCCTCGTGTGGGGAGAGACTGATTCGGGGTGACAGACACCTTATGGCCCG 743
Oy 292 GUAUAlaMetIlePheArgSerGlyProArgProAlaLeuArgSerasp---Serasp 310
Db 744 GACTTACATCTTTCGAAACCTTGGTCAGAAATCGAGTCTCGGAACAGAGCCCAATGAT 803
Oy 311 GlnSerIleuHisAspProArgPheValMetAlaAlaArgIleProGluAsnSerasp 330
Db 804 TCCCGCTGGCTCAATGAAACCAAGTTTGCAAGTCTTTGGATCCCAAGAGAGTGAAC 863
Oy 331 GlnAspAsnAspLysValTyrPhePheSerGluThrValProSerProAspGlyGly 350
Db 864 CCGATGACGATAAATCTATTTCTTCTCCGAGCTCCGAGTCTTCCGAGAGCAGCA 923
Oy 351 SerAsnHisValThrValSerArgValGlyArgValCysValAsnAspAlaGlyGly 370
Db 924 ATGGGGCGCATGTGTGTCTGTGGCCAGATCTGCAGATGACCTGGGGGCGCAG 983
Oy 371 ArgValIleuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValPro 390
Db 984 CGGAGCTGGTGCACAAATGACACCAATTTCTGAAGCGCGGCTTGTGTCTCAGTACT 1043
Oy 391 GlyProGlyGlyAlaGluThrHisPheAspGlnLeuAspValPheLeuLeuTrpPro 410
Db 1044 GGAGTTGAGGGT---GACACCCACTTGTACCACTTCAGATGATTTTCTCTCTCTCC 1100
Oy 411 LysAlaIleGlyLysSerLeuGluValTyrAlaLeuPheSerThrValSerAlaValPheGln 430
Db 1101 CGAGACCGCCAGACACCTTCTATGCTGTCTTCCACCTCCAGAGTGTCTCCAG 1160
Oy 431 GlyPheAlaValCysValTyrHisMetAlaAspIleTrpGluValPheAsnGlyProPhe 450
Db 1161 GGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1220
Oy 451 AlaHisArgAspGlyProGlnHisGlnTrpGlyProTyrGlyGlyLysValProPhePro 470
Db 1221 GCTCACAAAGAGGGGCGCTACACACAGGCGGTCTACACAGGCGGTGTCTCCCA 1280
Oy 471 ArgProGlyValCysProSerLysMetThrAlaGlnProGlyArgProPheLysSerThr 490
Db 1281 AGACCTGGCATGTGCCCCGACAGACCTTTGGC-----ACCTTCAGCTCCACC 1328
Oy 491 LysAspTyrProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpPro 510
Db 1329 AAGGACTTCCCAATGACGTTATCCAGTTTCTCGGAAACACCTCTCATGTATACACCA 1388
Oy 511 ValArgProArgHisGlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHis 530
Db 1389 GTCCTGGCCATGGGGGGGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1448
Oy 531 GlnIleValValAspArgValGluAlaGluAspGlyThrTyrAspValIlePheLeuGly 550
Db 1449 CAATATCCCGCCAGACAGAGACGCGCGCATGACACACATGTTCTTCTCATTTGGT 1508
Oy 551 ThrAspSerGlySerValLeuLysValIleAlaLeuGlnAlaGlyLysSerAlaGluPro 570
Db 1509 ACAGATGTGGGACACAGTGTGAAGATGATCTGATCCCAAGAGCGCACCTTAATCT 1568
Oy 571 GluGluValValLeuGluGluLeuGlnValPheLysValProThrProIleThrGlnMet 590
Db 1569 GAAGAGCTTACCTGGAAGAGCTGACAGTGTGAGGACCTGCGCTATCACACAGATG 1628
Oy 591 GluIleSerValLysArgGlnMetLeuTyrValGlySerArgLeuGlyValAlaGlnLeu 610
Db 1629 CAATCTCTCTAAAG----- 1645
Oy 611 ArgLeuHisGlnCysGluThrTyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAsp 630
Db 1646 -----GGCGCGCGCTGCGAGATGCTGCTGCGCGCGAT 1681
Oy 631 ProTyrCysAlaTrpAspGlyAlaSerCysThrHisTyrArgProSerLeuGlyLysArg 650
|||||

```

```

Db 1682 CCTTACTGCGCGTGGGATGATGATCACTGTGCACACGCTTCCAGCCTACG---GCCAGAGA 1738
Oy 651 ArgPheArgArgGlnAspIleArgHisGlyAsnProAlaLeuGlnCysLeuGlnSer 670
Db 1739 CGGTTCGGAGAGCAGACATAGAAATGACGACCCAGACCCATGCTGTCTGAGACTCT 1798
Oy 671 GlnGluGluGluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGlnHisAsn 690
Db 1799 TCTCAC-----TCTGTGTCTGCGGAGAGAGAGTGTGGGTGTGAGAGCGGC 1846
Oy 691 SerThrPheLeuGluCysLeuProLysSerProGlnAlaAlaValArgTrpLeuGln 710
Db 1847 ACAGCGCTTCTGTGAGTGTGAGCCCGCTGCTCCAGCGCATGTGACATGACCTTCCAA 1906
Oy 711 ArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeuHisThrGlu 730
Db 1907 GGTGACGGGAGAGCAGCTACACACAGTGTGCGGTGAGAGAGAGTGAAGCGACTGCG 1966
Oy 731 ArgGlyLeuLeuPheArgArgLeuSerArgPheAspAlaGlyThrTyrThrCysThrThr 750
Db 1967 CGGGGCTGCTGTGGGGGCGTGGCGCGCCAGAGACTGTGGGTGTGTGCTTGGCTCGC 2026
Oy 751 LeuGluHisGlyPheSerGlnThrValArgLeuAlaLeuValIleValAlaSer 770
Db 2027 GTTGACACAGGCTTTCACACACACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2086
Oy 771 GlnLeuAspAsnLeuPheProProGluProLysProGluGluProAlaArgGlyGly 790
Db 2087 CAGGCTGACAGACTG-----GCACGGGACGAG 2113
Oy 791 LeuAlaSerThrPro-----ProLysAlaTrpTyrLysAspIleLeuGlnLeu 806
Db 2114 GAAACAGCCGCTCTGCACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2173
Oy 807 Ile-----GlyPheAlaAsnLeuProArgValAspGluTyrCysGluArg 821
Db 2174 GTGAGACGAGCGGTGGCGAGGTGCAAACTCCCTCGCAATG----- 2215
Oy 822 ValTrpCysArgGlyThrThr-----GluCysSerGlyCysPheArgSerArgGly 840
Db 2216 -----TCCCGCCCGCAGCGCGGACCCACTGTGTGCGACGACATTCAGTCGTAAGGT 2269
Oy 841 LysGlnAlaArg 844
Db 2270 CGCAACAGACGG 2281

```

RESULT 2
 BI067436
 LOCUS
 DEFINITION
 chicken
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 chicken
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE
 1 (bases 1 to 647)
 AUTHORS
 COBURN, L.A., MORGAN, R.W. and BURNSIDE, J.
 TITLE
 Chicken ESTs from fat
 JOURNAL
 Unpublished (2001)
 CONTACT
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: coburnudel.edu, www.chickest.udel.edu.
 FEATURES
 1..647
 source
 /organism="Gallus gallus"

03-05-813-290-2 (1-8/5) X BI067436 (1-647)

[illegible]

RESULT	3
AM803853/c	
LOCUS	AM803853
DEFINITION	H12-UMD002-240300-058-F08 UMD082 Homo sapiens CDNA, mRNA sequence.
ACCESSION	AM803853
VERSION	AM803853.1 GI:7855736
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

[illegible]

CONTACT
COMMENT

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP
Brazil
Tel.: +55-11-3704-0023

Fax: +85-11-24049822
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=tl2-0M0082-240300-05-F08&t3=2000-03-24&t4=1>
 Seq primer: puc 18 forward
 High quality sequence start: 389.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="UM0082"
/dev_stage="Adult"
/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 199,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      86 a      140 c      124 g      70 t
ORIGIN

```

Alignment Scores
Pred. No.:
Score:
Percent Similarity
Best Local Similarity
Query Match:
DB:

US-09-813-290-2 (1-875) x AW803853 (1-420)

QY	382	LysAlaArgLeuValCysSerValProGlyProGlyGlyAlaGluThrHisPheAspGln	401
Db	360	AAGCCAGGCGTGTCTGCTCGGTGCCGCGCTGTGTGTGGCGAGCCACTTTGGACGAG	301
QY	402	LeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSerLeuGluValIlyrAlaLeu	421
Db	300	CTAAGAGATGTGTTCCTGCTGTGGCTTAAGCGCGGGAAGAGCCTCGAGGTGTACCGGCTG	211
QY	422	PheSerThrValSerAlaValPheGlnGlyPheAlaValLysValIlyrHisMetAlaAsp	441
Db	240	TTTCAGCAACCGTCATGCCCGTGTCCAAAGGCGTTGTGCGCTGTGTGACCAACAAGCGCAGAC	181
QY	442	IleTrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGlnHisGlnTrpGly	461
Db	180	ATCTGGAGAGTTTTCACAGCGGCGCTTTGGCCACGAGATGGGGCTCAGGACCAAGTGGGGG	121

QY 462 ProTyrGlyGlyValPProPheProArgProGlyValAlcysProserLysMetThrAla 481
 DB 120 CCCTATGCGGGGAGAGGCGCTTCCCTCCCTGCGCTGTCGCCACCAAGATGACGCA 61
 QY 482 GluProGlyArgProPheGlySerThrLysAspTyrProAspGluValLeuGlnPheAla 501
 DB 60 CAGCCAGAGCGGCTTGTGGCAGCAGCAGACGACTACCATGATGATGTCGCACTTGGC 1
 RESULT 4
 BI652258 990 bp mRNA linear EST 12-SEP-2001
 LOCUS 603299702F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5340217 5'
 DEFINITION mRNA sequence.
 ACCESSION BI652258
 VERSION BI652258.1 GI:15566494
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-remail.nih.gov
 Tissue Procurement: Lohar Hennighausen Ph.D., Chu Xia Deng, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHM11864 row: C column: 02
 High quality sequence stop: 892.
 Location/Qualifiers
 1. 990
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_image="5340217"
 /clone_id="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; NCI;
 Site_2: Sali; Cloned unidirectionally. Primer: Oligo dr.
 Average insert 2 kb. Library constructed by Life
 Technologies, catalog #12017-018. Investigators providing
 samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
 for transgenic model; Xu et al., Nature Genetics 22, 37-43
 (1999). Note: this is a NCI_CGAP Library."
 BASE COUNT 194 a 285 c 292 g 219 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.45e-47 Length: 990
 Score: 777.50 Matches: 163
 Percent Similarity: 65.36% Conservative: 54
 Best Local Similarity: 49.10% Mismatches: 95
 Query Match: 16.38% Indels: 21
 DB: 10 Gaps: 6
 US-09-813-290-2 (1-875) x BI652258 (1-990)
 QY 184 ArgGluValLeuTyrProProGlnPro--GlyGln-ArgGluGluCysValArgLysG 202
 DB 40 AGACGGCTCTGTAACCGCTGAGCGCCCGCTGGAATGCGAATGCACTGGCGCAG 99
 QY 202 YArgAspProLeuThrGluCysAlaAsnPheValArgValLeuGlnProHisAsnArgTh 222
 DB 100 GAAGGCAATGTGACCGAGTGCATGAACCTCGTGAAGTGTGTGACGCGCTACAAACACAC 199
 QY 222 rHisLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuLeuThrValG 242

DB 160 CCACCTGTGGCGCTGTGGCAGAGCGCTTCCACCAACGCGCGCTTGTGGAGGTGG 219
 QY 242 YHisArg---GlyLumiValLeuHisLeuGlnProGlySerValGlySerGlyArgG 261
 DB 220 CCACCGGCTGGAGAGACCCATGCTTCAACCTGACCGGAGAACTTGGAGCGCAGAGG 279
 QY 261 YArgCysProHisGluProSerArgProPheAlaSerThrPheLeuAspGlyGluLeuTy 281
 DB 280 GAACACTCTTATGACCAACCAAGCGATCGGCGCTCGGTGTGTGGGGGAAGAGCTCTA 339
 QY 281 rThrGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetLeuPheArgSerGlyGlyPr 301
 DB 340 TTCTGGGTGGCAGCAGACGACTTATGAGCGGCGGACTTAACTTCTTCCAAACCTTGGCA 399
 QY 301 oArgProAlaLeuArgSerAsp---SerAspGlnSerLeuLeuHisAspProArgPheVa 320
 DB 400 GAATCCGAGTCTCCGACAGAGCCCAATGATTCGCCGTGCTCAATGACCAAGTTGT 459
 QY 320 lMetAlaAlaArgGlyLeuProGluAsnSerAspGlnAspAsnAspLysValTyrPhePhe 340
 DB 460 CAAGTCTTTGGATCCAGAGAGAGAGAGACCCGATGATGACGATTAATCTATTCTCTCT 519
 QY 340 eSerGluThrValProSerProAspGlyGlySerAsnHisValThrValSerArgValG 360
 DB 520 CCGGAGTCCCGCTGTGGAAGCAGCAGCAATGGGGCGCATGTGTGTCTGTGTGG 579
 QY 360 YArgValCysValAsnAspAlaGlyGlyArgValLeuValAsnLysTyrSerThrPh 380
 DB 580 CCAGATCTGCGAGATGACCTGGGTGGCGGAGCTTGTGTGCAACAAATGACCAACTT 639
 QY 380 eLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyAlaGluThrHisPheAs 400
 DB 640 TCTGAAGCGCGGCTGTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 696
 QY 400 pGlnLeuGluAspValPheLeuLeuTyrProLysAlaGlyLysSerLeuGluValTyrAl 420
 DB 697 CCACCTTGAAGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
 QY 420 aLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCysValTyrHisMet-A 440
 DB 757 TGTCTTCTCCACCTCCAGTGTCTCTCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGA 816
 QY 440 lAspPLeuTyrGluValPheAsnGlyProPheAlaHisArgAspGlyProGlnHisGlnT 460
 DB 817 CCGATGTGCGCGAGACCTTCTGCGGACCTATGAGCTCAACAGAGGCGCTTACACCACT 876
 QY 460 rGlyProTyrGlyGlyValPProPheProArgProGlyValCysProserLysMet 480
 DB 877 GGGTGTCTTACCGGGGTGGGTCTTATCCAGACCTGATGTGCGCCAGCAAGACTT 936
 QY 480 hAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAspGluValLeuGlnP 500
 DB 937 GGGCC-----CTTGGCTCCACAGAGGC-TTCCGAATGGCGTTCCCGCT- 979
 QY 500 heAlaArgAlaHisProLeuMetPheTyrPro 510
 DB 980 -----TGGCGG 985
 RESULT 5
 BI689389 785 bp mRNA linear EST 18-SEP-2001
 LOCUS 603315834F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:535609 5'
 DEFINITION mRNA sequence.
 ACCESSION BI689389
 VERSION BI689389.1 GI:15652018
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9qabp-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1904 row: d column: 10
High quality sequence stop: 783.

FEATURES

source
Location/Qualifiers
1. 785
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5355609"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies, Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 161 a 239 c 200 g 185 t

ORIGIN

Alignment Scores:
Pred. No.: 6.09e-45 Length: 785
Score: 749.50 Matches: 152
Percent Similarity: 71.59% Conservative: 37
Best Local Similarity: 57.58% Mismatches: 68
Query Match: 15.79% Indels: 7
DB: 10 Gaps: 2

US-09-813-290-2 (1-875) x B1689389 (1-785)

QY 354 ValThrValSerArgValGlyArgValCys-ValAsnAspAlaGlyGlyGlnArgValLe 373
Db 5 ATGCTGTGTCTGCTGTTGGCCACATCTGCGAGAGAGTACCTGGGTGGCCAGGAGCTT 64
QY 373 uValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSerValProGlyProG 393
Db 65 GGTCAACAATGAGACCACTTGTGAAGGCCGGCTGTGTGCTCAGTACCTGGAGTTGA 124
QY 393 YGLValAGLThrHisPheAspGlnLeuGluAspValPheLeuLeuTrpPolysAlaG 413
Db 125 GGGT---GACACCCAGCTTGTGACCACTTCAGATGTTTCTTGTCTCCCGAGACCG 181
QY 413 YLysSerLeuGluValTyraLeuAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAl 433
Db 182 CCAAGACACCTTCTCATGTCTGCTTCTCCACCTCAGTGTCTTCCAGGGCTCTGC 241
QY 433 aValCysValTyHisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisAr 453
Db 242 TGTGTGTGTGTACAGCATGTGACGATGTGCGCCAGCCTTCTTGGACCTTTTGTCTCA 301
QY 453 GAAPGlyProGlnHisGlnTrpGlyProTyrglyGlyValValProPheProAlaProG 473
Db 302 AGAGGGGCTTACACACAGTGGGTGTCTTACAGAGGTGTGTCTTCCCAAGACCTGG 361
QY 473 yValCysProSerLysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTy 493
Db 362 CATGTGCCCGACAGACACTTGGC-----ACCTTACGCTCCACCAAGAGACTT 409
QY 493 rProAspGluValLeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValaLarPr 513
Db 410 CCCAGATGAGCTTATCCAGTTTGTCTGGAACACACCTCTCATGTACCAACCAAGTCTGCC 469

QY 513 oArghHisGlyArgProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleVa 533
Db 470 CATGGGGGGGCGCCCTCTCTTCTACAAAGTGGAGCTGGGTACACCTTCACCCAAATCCG 529
QY 533 lValAspArgValGluAlaGluAspGlyThrTyraSpValIlePheLeuGlyThrAspSe 553
Db 530 CCGAGACCGAGTACAGCGCTGCCGATGACACATGATGTTCTTCTTATTTGATGATGAT 589
QY 553 rGlySerValLeuLysValIleAlaLeuGlnAlaGlyGlySerAlaGluProGluGluVa 573
Db 590 GGGCACAGTGTGAAAGTATCTCAGTCCCAAGGACCGACCTAATTCATGAGGACT 649
QY 573 lValLeuGluGluLeuGlnValPheLysValProThrProIleThrGluMetGluIleSe 593
Db 650 TCTCTCGAAGAGCTGACAGGTGTGAGGAGCTGTGCCCTATCACCACGATGCAATATC 709
QY 593 rValLysArgGlnMetLeuTrVal-GlySerArgLeuGlyValAlaGlnLeuArgLeuH 613
Db 710 CTCTAAAGCAACAACACTACGTACGTACGATCCGACGCGACGATGCGCCAGATTGCTTGC 769
QY 613 lSgLnCys 615
Db 770 ATCGCTGC 777

RESULT 6 645 bp mRNA linear EST 15-JUN-2001
B1066459
LOCUS B1066459

DEFINITION pGln.pk008.d21 normalized chicken fat cDNA library Gallus gallus
CDNA clone pGln.pk008.d21 5' similar to g119910362
ref|NP_064548.1| semaphorin sem2 [Homo sapiens] g111449696
ref|XP_003233.1| semaphorin sem2 [Homo sapiens] dbj|BA8132.1|
(AB029496) semaphorin sem2 [Homo sapiens]g, mRNA sequence.
B1066459
B1066459.1 GI:14473981
EST.

VERSION B1066459.1
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 645)
AUTHORS Cogburn,L.A., Morgan,R.W. and Burnside,J.
TITLE Chicken ESTs from fat
JOURNAL Unpublished (2001)
COMMENT Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
source
Location/Qualifiers
1. 645
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pGln.pk008.d21"
/clone_lib="normalized chicken fat cDNA library"
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMDH10B"
/note="Vector: pSPORT1"

BASE COUNT 98 a 221 c 202 g 118 t 6 others

ORIGIN

Alignment Scores:
Pred. No.: 3.31e-42 Length: 645
Score: 710.50 Matches: 137
Percent Similarity: 74.30% Conservative: 22
Best Local Similarity: 64.02% Mismatches: 44
Query Match: 14.97% Indels: 11
DB: 10 Gaps: 4

US-09-813-290-2 (1-875) x B1066459 (1-645)

QY	98	Ala	Trp	---	Ala	Ile	Cys	Trp	Phe	Leu	Gly	Ile	Val	Leu	His	Gly	Ser	Ser	Gly	116		
Db	6	GGCGGCGACCGCTGGGGCTGGCTGGGGCGCGCTGGT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----		
QY	117	Pro	Ser	Pro	Gly	Pro	Ser	Val	Pro	Ala	Ile	Arg	Leu	Ser	Val	Arg	Asp	Phe	Leu	Ser	Ala	
Db	54	AGCAGCCCG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----		
QY	137	Asn	Arg	Ser	Ala	Ile	Phe	Leu	Gly	Pro	Gly	Ile	Ser	Leu	Ser	Leu	Asn	Gln	Ala	Met	Tyr	Leu
Db	102	AACCGCTCCGCTCTTTCTTTCTTTGGCGACCGGGCGCTTCCGTGGCGTTCCGCTCCCTCTACTCCG	161																			
QY	157	Asp	Gly	Tyr	Arg	Asp	Arg	Phe	Gly	Leu	Gly	Ile	Val	Asp	Ala	Ile	Tyr	Ser	Leu	Arg	Leu	
Db	162	GATGAGTACCGCTACCGGGCTTTCATTGGGGGCAAGATGTCCTTACTCCCTCGCTCG	221																			
QY	177	Asp	Gln	Ala	Trp	Pro	Asp	Pro	Arg	Glu	Val	Leu	Tyr	Pro	Pro	Gln	Pro	Gly	Gln	Arg	Glu	
Db	222	GACAGCGCCAGCGCCGACACCAAGAAATCTTACGCGCGCCACACGCTCGCTGGACACAGCGAG	281																			
QY	197	Gln	Cys	Val	Arg	Gly	Ser	Gly	Asp	Pro	Leu	Thr	Gly	Cys	Ala	Asn	Phe	Val	Arg	Val	Leu	
Db	282	GAGGTGTTTGGGAGGGGAGAGCCACACACGACTGTGCGCACTGTCCACTTACGTCCGCTCGT	341																			
QY	217	Gln	Pro	His	Asn	Arg	Tyr	His	Leu	Leu	Ala	Cys	Gly	Thr	Gly	Ala	Phe	Gln	Pro	Thr	Cys	
Db	342	CAGCGCTACACACAGACACACCTGTGGCTGTGGGAGGGGTGCTGCACCCGCTGTC	401																			
QY	237	Ala	Leu	Ile	Thr	Val	Ala	Gly	His	Arg	Gly	Leu	-----	His	Val	Leu	His	Leu	Gln	Pro	Gly	
Db	402	ACCTTCACTTACGTTGGGGGACCGGGGTGAAGACCAACGATGCTTCAAGTTGATGACCTGTGT	461																			
QY	255	Ser	Val	Glu	Ser	Gly	Arg	Gly	Ala	Cys	Pro	His	Glu	Pro	Ser	Arg	Pro	Phe	Ala	Ser	Thr	
Db	462	AGCAGGAGAGTGTGTCGTGGCAGGTGGCCACATGAGCTGACCGTCGATTTGGCCAGCAC	521																			
QY	275	Phe	Ile	Asp	Gly	Gly	Leu	Tyr	Thr	Gly	Leu	Thr	Ala	Asp	Phe	Leu	Gly	Ala	Arg	Gln	Ala	Met

[illegible]

RESULT 8
AL584625 754 bp mRNA linear EST 28-FEB-2001
LOCUS AL584625 Stratagene Chick Embryo Lambda cDNA library (* 937405)
DEFINITION Gallus gallus cDNA clone ROS012B12, mRNA sequence.
ACCESSION AL584625
VERSION AL584625.1 GI:13163356
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 754)
REFERENCE Murray, F.
Stratagene Chick Embryo Lambda cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Frazer Murray
Dept. Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UK
Tel: +44 (0)131 527 4200
Fax: +44 (0)131 440 0434
Email: frazer.murray@bbsrc.ac.uk
Seq primer: T3.
Location/Qualifiers
1..754
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone_lib="Stratagene Chick Embryo Lambda cDNA library (* 937405)"
/tissue_type="Embryo"
/dev_stage="5 days old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="vector: pBLUESCRIPT SK; Site-1: EcoRI; Site-2: XhoI
; Cloned unidirectionally. Primer: Oligo dT, Uni-ZAP XR
vector. Average insert size: 1.5kb; 5' adaptor sequence:
5' GAATTCGGCAGCAG 3'; 3' adaptor sequence: 5'
CTCCGATTTTTTTTTTTTTTTT 3'."

BASE COUNT 158 a 242 c 218 g 134 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 8,36e-41 Length: 754
Score: 692.50 Matches: 138
Percent Similarity: 68.80% Conservative: 34
Best Local Similarity: 55.20% Mismatches: 66
Query Match: 14,598 Indels: 13
DB: 9 Gaps: 5

US-09-813-290-2 (1-875) x AL584625 (1-754)

QY 423 SerThValSerAlaValPheGlnGlyPheAlaValCysValTyrHisMetAlaAspIle 442
DB 15 TCCGCGATCGGGGTGGGTGTTCAAGGCTCGGCTGCTGCTTACCTCCAGCGCGACATC 74
QY 443 TrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGlnHisGlnTrpGlyPro 462
DB 75 CCGATGGTGTCAATGAGGCGCTCGCGCACAAAGAGAGGCCCAACACACAGTGGATGCC 134
QY 463 TyrGlyGlyValPheProPheProArgProGlyValCysProSerLysMetThrAlaGln 482
DB 135 TACAGGGGAAATACCAACCCCGCGCGACCTGCCCGGGGG--ACCTTCACA 191
QY 483 ProGlyArgProPheGlySerThrLysAspTyrProAspGluValLeuGlnPheAlaArg 502
DB 192 CCCTCC-----ATGAAGTCAACCAAGAGACTACCCGAGCAAGTGTATCAACTTCATGCCG 245
QY 503 AlaHisProLeuMetPheThrProValArgProArgHisGlyArgProValLeuValLys 522
DB 246 TCCACACCTCTGTATGTACACAGCGGCTACCGCGGCCACCGGACGCTCTGCTGCTGCCG 305

QY 523 ThrHisLeuAlaGlnGluLeuHisGlnIleValValAspArgValGluAlaGlyAspGly 542
DB 306 ACCAAGCTCACTACCGCTTACACCACTCGCTGTGACACAGGTGACGCGGACGATCGG 365
QY 543 ThrTyrAspValIlePheLeuGlyThrAspSerGlySerValLeuValIleAlaLeu 562
DB 366 CGCTATGAGTGTCTTTCTCTGGCAGATCGGGCAGATGAGCAAGATGATGTGCTG 425
QY 563 GlnAlaGlyCysSerAlaGluProGluGluValValLeuGluGluLeuValPheLys 582
DB 426 CCCCCG---GATGACATGTAGACAGACAGAGAGAGTACGTGAGAGATGACAGTGTAC 482
QY 583 ValProThrProIleThrIleGluMetGluIleSerValLysArgGlnMetLeuTyrValGly 602
DB 483 GTGCCAGCACCCTATCAAGATGATGACCATCTCCCTCCACAGAGCAACACTTACGTGCC 542
QY 603 SerArgLeuGlyValAlaGlnLeuArgPheHisGlnCysGluThrTyrGlyThrAlaCys 622
DB 543 TCGGACAGTAGAGTAGACCCACCTGCGCTGCACCGGTGTGACGTGTATGAGAAAGCTGT 602
QY 623 AlaGlyCysCysLeuAlaArgAspProTyrCysAlaThrAspGlyAlaSerCysThrHis 642
DB 603 GCTACTGCTGCT--GCCCGGACCCACTACTGTGCTGCGATGCGAGGCGCTGCAGCGCC 660
QY 643 TyrArgProSerLeuGlyLysArgArgPheArgGlnAspIleArgHisGlyAsnPro 662
DB 661 TAC-----TCACATCTCTCAAGCGGACAGCGGAGNAGACGTCCGACGACGACCC 714
QY 663 -----AlaLeuGlnCys 666
DB 715 ATGCGCGGTGCGGAGTAGACACTTCATATGC 744

RESULT 9
BI067451 576 bp mRNA linear EST 15-JUN-2001
LOCUS BI067451
DEFINITION pGln.pK012.d9 normalized chicken fat cDNA library Gallus gallus
cDNA clone pGln.pK012.d9 5' similar to gb|AA86899.1 (U28243)
collapsin-5 [Gallus gallus]G, mRNA sequence.
ACCESSION BI067451
VERSION BI067451.1 GI:14474973
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 576)
REFERENCE Cogburn, L.A., Morgan, R.W. and Burnside, J.
Chicken ESTs from fat
Unpublished (2001)
COMMENT Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.
Location/Qualifiers
1..576
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone_lib="pGln.pK012.d9"
/clone_lib="normalized chicken fat cDNA library"
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMD10B"
/note="Vector: pSPORT1"

BASE COUNT 106 a 179 c 161 g 104 t 26 others
ORIGIN

Alignment Scores:
Pred. No.: 8,96e-41 Length: 576
Score: 690.00 Matches: 135

```
Conservative: 11
Mismatch: 40
Indels: 1
Gaps: 0
```

FEATURES		High quality sequence stop: 803.
source	Location/Qualifiers	
	1..823	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:5315376"	
	/clone_1id="NIH_CGAP_Mam3"	
	/tissue_type="tumor, gross tissue"	
	/lab_host="DH10B"	
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; NCI; Ate-2; Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI CGAP Library."	
BASE COUNT	166 a 235 c 232 g 190 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	3.04e-40	Length: 823
Score:	685.50	Matches: 144
Percent Similarity:	68.73%	Conservative: 45
Best Local Similarity:	52.36%	Mismatches: 82
Query Match:	14.44%	Indels: 6
DB:	10	Gaps: 3
US-09-813-290-2 (1-875) x B1645606 (1-823)		
QY	213 VALATGVALLeuGlnProHisAsnArgPrrHisLeuAlaCysGlyThrGlyAlaPhe	232
Db	1 GTGAAGCTCTACAGCCCTACAAACACACCCACTTCTGCGCTGTGGCACAGGGCTTTC	60
QY	233 GlnProThrCysAlaLeuLeuThrValAlaGlyHisArg--GlyGluHisValLeuHisLeu	251
Db	61 CACCCACCTGGCGCTTGTGTGAGGTGGGGCCACCGGCTGAGAGAACCCATGCTTCAACTG	120
QY	252 GluProGlySerValGluSerGlyArgGlyArgGysProHisGluProSerArgProPhe	271
Db	121 GACCGGAGGAACCTTGAGGAGCGCAAGGGGGAAGACCTCTATGACCCAGGATCGGGCT	180
QY	272 AlaSerThrPheIleAspArgGlyGluLeuGlyTrpGlyLeuThrAlaAspPheLeuGlyArg	291
Db	181 GCCTCGTGCTGTGGGGGGAAGAGCTGTATTCGGGGGTGGACACACACCTTATGGGCGG	240
QY	292 Glu-AlaMetIlePheArgSerGlyGlyProArgProAlaLeuArgSerAsp--SerAs	310
Db	241 GACTTACCACATCTTTCGAAAGCCTTGTCAGAACTCCAGACTCCGAACAGCCCATGA	300
QY	310 PoinSerLeuLeuHisAspProArgPheValIleMetAlaAlaArgL1ProGluAsnSerAs	330
Db	301 TTCCCGCTGGCTCAATGAAACCCAGATTGTGTCAAGAGCTTTGGATGCCAGAGGTAGAA	360
QY	330 PGLnAspAsnAspIysValTyrTrpPhePheSerGluTrValProSerProAspGlyG1	350
Db	361 CCCTGATGACGATAAATATATATTCTTCCCGCCAGCTCGGTGGAAAGCACACACAGC	420
QY	350 YSerAsnHisValThrValSerArgValGlyArgValCysValAsnAspArgIaGlyGlyG1	370
Db	421 AATGGGCGGCATGTCTGTGTCTGTGTGGCCAGATCTGAGGAATGACCTGGGTGGCCA	480
QY	370 nArgValLeuValAsnIysTrpSerTrpPheLeuIlySAlaArgLeuValCysSerValPr	390
Db	481 GCGGAGCTTGTCACAAATAGGACCAATTTCTGAAAGCGCGGCTGTGTGCTCAGTACC	540
QY	390 OG1YProGlyGlyAlaGluThrHisPheAspGlnLeuGluAspValPheLeuLeuTrpPr	410
Db	541 TGGAGTTGAAGG--GACACCACTTTGACCACTTCAGAGATGTTTTCTTCTGTCTC	557
QY	410 OLYSAlaGlyLysSerLeuGluValTyrAlaLeuPheSerThrValSerAlaValPheG1	430
Db	598 CCGAGACCGGACACACCTTGTCTATATGCTGTCTTCTCCACCTCCAGTGGTGTCTTCA	657

AUTHORS Murray, F.
 TITLE Stratagene Chick Embryo Lambda cDNA Library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Frazer Murray
 Dept: Genomics and Bioinformatics
 Roslin Institute
 Roslin, Midlothian, EH25 9PS, UK
 Tel: +44 (0)131 527 4200
 Fax: +44 (0)131 440 0434
 Email: frazer.murray@bbsrc.ac.uk
 Seq primer: T3.

FEATURES
 source Location/Qualifiers
 1. 786
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="ROS007F06"
 /clone_lib="Stratagene Chick Embryo Lambda cDNA Library (*
 937405)"
 /issue_type="Embryo"
 /dev_stage="5 days old"
 /note="Vector: pBLUESCRIPT SK⁺ Site: 1: EcoRI, Site 2: XhoI
 ; Cloned unidirectionally. Primer: Oligo dT, Uni-ZAP XR
 vector. Average insert size: 1.5kb.; 5' adaptor sequence:
 5' GAATTCGCGACGAG 3' / 3' adaptor sequence: 5'
 CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT 166 a 242 c 218 g 155 t 5 others
 ORIGIN

Alignment Scores:
 Pred. No.: 8.97e-39 Length: 786
 Score: 665.00 Matches: 133
 Percent Similarity: 71.02% Conservative: 41
 Best Local Similarity: 54.29% Mismatches: 65
 Query Match: 14.01% Indels: 7
 DB: 9 Gaps: 4

US-09-813-290-2 (1-875) x AL584324 (1-786)

QY 322 AAlAlAlArllePrrogluAnserAspGlnaSpasnsplysValTyrPhePheSer 341
 Db 5 TCGGCAGAGCCCTGCTCCGTCGAGAGATGATGACAAACTCTACTTCTCTCGG 64
 QY 342 GluThrValProserProAspGlyGlySerAsnHisValThrValSerArgValGlyArg 361
 Db 65 GAG-----AAGTGGGGGAGATGCCCGCTCAGCCCGGGGCTATTCGGCATGGGGGC 118
 QY 362 ValCysValasnaSpalaglyGlnArgValLeuValAsnLysTTPSerThrPheLeu 381
 Db 119 ATGTGCTGAAGCAGATGAGACACTGCTGCTGGTGAACAATGAGACACTTCTCTG 178
 QY 382 LysAlaArgLeuValCysSerValProGlyProGlyGlyGluThrHisPheAspGln 401
 Db 179 AAGCCCGGCGTCTCTCCGTCGAGACCCGATGGGATCGAGACACTTGTGATGAG 438
 QY 402 LeuGluAspValPheLeuLeuThrProLysAlaGlySerLeuGluValTyrAlaLeu 421
 Db 239 CTTAGAGAGCTCTTCATCCAGACAGACAGACACCAAGATCCGTATCTAGCCGCTC 298
 QY 422 PheSerThrValSerAlaValPheGlnGlyPheAlaValCysValTyrHisMetAlaASP 441
 Db 299 TTTCTCCGATCGGGGCTGCTTCAGGGCTGCTCTCTGCTCTGCTCATGCGCCAGC 358
 QY 442 LLeTrrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGlnHisGlnTrrpGly 461
 Db 359 ATCCGCAATGTTTCATATGGCCCTTCGCGACAGAGGGGTCCCAACTACAGTGGATG 418
 QY 462 ProTrrpGlyGlyValProPheProArgProGlyValCysProSerLysMetThrAla 481
 Db 419 CCTACAGCGGCAAAATACCTTACCCCGGGCGGACCTCGCCCGGGGGG--ACCTTC 475
 QY 482 GlnProGlyArgProPheGlySerThrLysAspTrrProAspGluValLeuGlnPheAla 501
 |||

Db 476 ACAACCTCC-----ATGAGTCACACCAAGCACTACCCGACGAAGTATCAACTTCATG 529
 QY 502 ArgAlaHisProLeuMetPheTrrProValArgProArgHisGlyArgProValLeuVal 521
 Db 530 CGCTTCACCTCTGTATGTATACAGCGCTGTACCGCGCCACGCGGAGCTCTGTGTG 589
 QY 522 LysThrHisLeuAlaGlnGlnLeuHisGlnIleValValAspArgValGluAlaGluASP 541
 Db 590 CGACCAAGCAAGTACAGTACCGCTTAC--ACCATCGCTGTGACAGTGTGAGCGGACAGAT 648
 QY 542 GlyThrTyrAspValIlePheLeuGlyThrAspSerGly---SerValLeuLysValIle 560
 Db 649 GCGCGCTATGAGTCTTTTCTGCGCAAAATCGGGGACGCGTGCAGAAAGTCATTC 708
 QY 561 AlaleuGlnAlaGly 565
 Db 709 TGCTTGACCCCGGGA 723

RESULT 13
 B1646227
 LOCUS B1646227 737 bp mRNA linear EST 12-SEP-2001
 DEFINITION 603276957F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5317143 5',
 mRNA sequence.
 ACCESSION B1646227
 VERSION B1646227.1 GI:15560463
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS NIH-MGC
 TITLE NIH-MGC http://mgc.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: ggaups@mail.nih.gov
 Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: L16M1804 row: a column: 16
 High quality sequence stop: 724.

FEATURES
 source Location/Qualifiers
 1. 737
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5317143"
 /clone_lib="NIH_CGAP_Mam3"
 /issue_type="tumor, gross tissue"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert 2 kb. Library constructed by Life
 Technologies, catalog #12017-018. Investigators providing
 samples: Lotmar Hennighausen/Chu-Xia Deng, NIH Reference
 for transgenic model: Xu et al., Nature Genetics 22, 37-43
 (1999). Note: this is a NCI_CGAP Library."

BASE COUNT 157 a 226 c 195 g 159 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.08e-38 Length: 737
 Score: 659.50 Matches: 141
 Percent Similarity: 67.84% Conservative: 32
 Best Local Similarity: 55.29% Mismatches: 63
 Query Match: 13.90% Indels: 21
 DB: 10 Gaps: 4

US-09-813-290-2 (1-875) x B1646227 (1-737)

LOCUS	GI:8184444	778 bp	mRNA	linear	EST 21-APR-2003
DEFINITION	BG184444 Athensys RAGE Library Homo sapiens cDNA, mRNA sequence.				
ACCESSION	BG184444				
VERSION	BG184444.1 GI:13706131				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 778) Harrington,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozar,S., Mays,R., Smith,E., Veloso,N., Kikka,A., Hess,J., Cothen,K., Lo,K., Offenbacher,J., Dandig,J. and Ducar,M.				
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)				
MEDLINE	21227151				
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scjain@atersys.com High quality sequence stop: 550. Location/Qualifiers 1..778 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Athersys RAGE Library" /cell_line="HT1080" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."				
FEATURES	source				
BASE COUNT	224 a 165 c 187 g 201 t 1 others				
ORIGIN					
Alignment Scores:					
Pred. No.:	7,23e-38	Length:	778		
Score:	652.50	Matches:	133		
Percent Similarity:	68.06%	Conservative:	46		
Best Local Similarity:	50.57%	Mismatches:	71		
Query Match:	13.75%	Indels:	14		
DB:	10	Gaps:	4		
US-09-813-290-2 (1-875) x BG184444 (1-778)					
Qy 248	ValLeuHisLeuGluuProGlySerValGluSerGlyArgGlyArgGlyProHisGluPro	267			
Db	9 CAGCTGCACCTGGTATACACCACAGATGTGAAGAGAGAGGCGCAGATGCTTTTGACCCC	68			
Qy 268	SeArgProPheAlaSerThrPheIleAspGlyGluLeuGlyThrGlyLeuThrAlaAsp	287			
Db	69 AGCTCTCTCTTCATCTCCACTTAATGTAGTAGAATGTGTTGGTGACCTACAGTCAAC	128			
Qy 288	PheLeuGlyArgGluAlaMetIlePheArgSerGlyLysProArgProAlaLeuArgSer	307			
Db	129 TACTGGAGCGAGAACCCCTGATCTCCGAGATGGGGCGACTGCCCATATCCGACT	188			
Qy 308	Asp---SerArgInseTleuLeuHisAspProArgPheValMetAlaIleArgIlePro	326			
Db	189 GACATGACATGACATGACGCTCTGTGAAGAAGACCAAACTGTAGGTGATACATGATTCCT	248			
Qy 327	GluAsnSerArgPheAlaSerAspAsnAspLysValTyrPhePhePheSerGluThrValProSer	346			
Db	249 GACAAATGAGACAGACATGACCAACAAAGTATATTCTTTTACTGAGAGGCACTGGAG	308			
Qy 347	ProAspGlyLysLeuSerHisValThrValSerArgValGlyArgValLysValAsnAsp	366			

Db	309	GCAGAAACAATGCT--CACGCAATTACACCAGGAGTGGCGACTGTGTGAATGAT	365
QY	367	AlAGlYglYnIarYvalLeuValaAsLysTrpSerThPheLeuYsAlaArgLeuVal	386
Db	366	GTAGGAGGCGAGAAATACGTGTGAATTAATGTGAGACCTTCTAAAGCGAAGCTGTT	425
QY	387	CysSerAlrProglYrProglYsIglValIgluThrHisPheaspInleuGluaSvpAlaPhe	406
Db	426	TGTCATATCCAGGAATGAATGGAAATGACACATATTATTATGAATATAGAGACGTTTTT	485
QY	407	LeuLeutrrProYsAlaGlyLysSerLeuGlValAlrYalAlaPheSerThrValSer	426
Db	486	TGTGTACTACACAGATCATAGAATCCAGTGATATTGTGACTCTTATACACTACAGAT	545
QY	427	AlaValPheGlnGlyPheAlaValCysValIlyrHisMetLaaspIleTrpGluValPhe	446
Db	546	AATATTCTTTCAGGCGCAT-GCTAAATGTGTCTATCCACTGTCTAGCATTCGGGACGCTTC	604
QY	447	AsnGlnrProPheAlaHisIAsrYaspGlyProGlnHisGlnTrpGlyProYrGlyGlyLys	466
Db	605	AACGACCATATGCACTATAGGAAGAGCCTGGAATACCATCGTGTCAGTCTTGAAGGAAAA	664
QY	467	ValrProPheProArGTrpGlyValCysrProSerLysMetThrAlaGlnProglYarYrPro	486
Db	665	GTCCCTTATCCAAAGCGCTGGTCTTGTGGCCAGCAAGNAAT-----GGAGGAGAG	715
QY	487	PheGly-SerThrLysAspTyrProAspGluValLeuGlnPheAlaArgAlaHisProL	506
Db	716	ATAGGAAACCAACANGACTATCTCT-----TGAAGATGCCATCGCA	757
QY	506	eumet 507	
Db	758	TGTTG 762	

```
Search completed: October 9, 2002, 21:32:13
Job time : 1657 secs
```

